GEOMETRIC MORPHOMETRIC COMPARISON of Cranium of PRIMATES along the evolutionary pathway

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT

FOR THE DEGREE OF

Bachelor of Technology
Biomedical

By

Aniket Mazumdar

110BM0635

Department of Biotechnology and Medical Technology
National Institute of Technology Rourkela
2014

GEOMETRIC MORPHOMETRIC COMPARISON of Cranium of PRIMATES along the evolutionary pathway

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT

FOR THE DEGREE OF

Bachelor of Technology
Biomedical

By

Aniket Mazumdar

Under the guidance of

Prof. Sirsendu Sekhar Ray

Assistant Professor

Department of Biotechnology and Medical Engineering

Department of Biotechnology and Medical Technology
National Institute of Technology Rourkela
2014

National Institute of Technology Rourkela

CERTIFICATE

This is to certify that the thesis entitled, **"GEOMETRIC MORPHOMETRIC COMPARISON of Cranium of PRIMATES along the evolutionary pathway"** submitted by **Mr. Aniket Mazumdar** in partial fulfillment of the requirements for the award of the Bachelor of Technology in Biotechnology and Medical Engineering with specialization in "Biomedical" at National Institute of Technology, Rourkela is an authentic work carried out by him under my supervision and guidance.

To the best of my knowledge, the matter embodied in this thesis has not been submitted to any other University/ Institute for the award of any other Degree or Diploma.

Prof. Sirsendu Sekhar Ray
Assistant Professor
Dept. of Biotechnology and Medical Engineering
National Institute of Technology Rourkela
Rourkela-769008

Declaration

I, **Aniket Mazumdar**, hereby declare that the thesis named **"GEOMETRIC MORPHOMETRIC COMPARISON of Cranium of PRIMATES along the evolutionary pathway"** is my original research work. This work includes the valuable contribution of others and every effort is being highlighted with due reference of literature and acknowledgment of collaborative research and discussions.

This work was done under the guidance of **Dr. Sirsendu Sekhar Ray**, Dept. of Biotechnology and medical Engineering, NIT Rourkela.

Aniket Mazumdar

Dept. of Biotechnology and Medical Engineering

National Institute of Technology Rourkela

Rourkela- 769008, Odisha

India

E-mail: aniket.mazumdar.11192@gmail.com

Mob: +91-8280103521

Date: 12.05.2014

Table of contents

Acknowledgment

Though only my name appears on the cover of this dissertation, a great many people have contributed to its production. I owe my gratitude to all those people who have made this dissertation possible and because of whom my experience here has been one that I will cherish forever.

My deepest gratitude is to my supervisor Prof. Sirsendu Sekhar Ray. I have been amazingly fortunate to have an advisor who gave me the freedom to explore on my own, and at the same time the guidance to recover when my steps faltered. He taught me how to question thoughts and express ideas. His patience and support helped me overcome many crunch situations and finish this dissertation.

A special thanks to my family, honorable teachers and all my dear friends for their love and concrete support without which this work would never have seen the light of the day.

Aniket Mazumdar

Dept. of Biotechnology and Medical Engineering

National Institute of Technology Rourkela

Rourkela- 769008, Odisha, India

E-mail: aniket.mazumdar.11192@gmail.com

Mob: +91-8280103521 Date: 12.05.2014

Abstract

The terminology of Morphometry had been initiated for the investigation of fossils of humans and animals. The actual preliminary idea itself was to appraise the similarities amongst the body parts to establish the evolutionary familiarity among different species.

Geometric Morphometrics is the factual examination of structure dependent upon Cartesian coordinates. Following the differentiation of shape from general size, position, and introduction of the landmark designs, the ensuing Procrustes shape directions might be utilized for measurable examination. Kendall shape space, the scientific space incited by the shape directions, is a metric space that could be estimated, principally by a Euclidean digression space. Therefore, it is necessary to effectively asses the similarity or variation within the entity form, shape, length of feature, as well as try to lay a trace on the process of evolutionary development. Importance should be on the process to undertake these investigations within this specific space of analysis. Deductions obtained from working systems like the Principle component analysis, multivariate reversion and partial least squares investigation are generally concentrated on the real shape or ideal deformations. These process in general save the above mentioned valuable properties. Asymmetry, reciprocal symmetry and lopsidedness in a sample can be obtained by Procrustes refinement along with its tagged replicate. Structure space can be protracted out of the morph space, dedicated to the analysis of structure independent contour. A general logarithm of the centroid dimension can be used to develop the shape coordinates and bring about the alteration. Centroid shape is an important measure in GM analysis, generally kept unassociated with shape variations so as to induce minimum isotropic discrepancy, when the area of interest undergoes any kind of

distortion. The addition expansion limit of the thin plate spline analysis is one of the most effective system to record deformation in lattices, as well as to undertake any kind of observance in three dimensional structures. This process is of equal importance when an analysis needs to be undertaken with missing landmarks of damaged and distorted data sets, as well as of effective use in partial landmark based calculations. These partial landmark calculations enhances the feasibility to integrate designs, schemes, layouts, planes and surfaces in the wake of any geometric morphometric research. The proficient optical and imagery visualization apparatuses employed along with the usually considerable measure of shape variables offer a means of upsurge to a specific exploratory style of examination, allowing the recognizing proof and quantification of previously equivocal and ambiguous shape attributes.

The process of Geometric Morphometric analysis has been used effectively for several years to establish the evolutionary hierarchy among various taxon of different era, various indigenous areas, location and among versatile populace. Diverse research investigation have been undertaken to lay emphasis on the effect of various environmental and ecological influences on this kind of variation and the subsequent process of evolution. In this present work, we have compared various primates such a Baboon, Gorilla, Chimpanzees, Homo sapiens, and tried to ascertain a liaison among them rendering to their shape and the constituent development over the evolutionary Hierarchy. The samples *0 (Baboon)*, 1 (*Chimpanzee*), 4 (*Gorilla*) and 7 (*Orangutan*) are morphometrically grouped together, while 6 (*Lesser Bushbaby*), 8 (*Ruffed Lemur*) and 10 (*Tarsier*) form a group of closely related species and the “rodent” like primates of 2 (*Common Marmoset*) and 9 (*Squirrel Monkey*) are in close proximity to each other. Homo *sapiens* sample is different from all other species due to variation in characteristics, and standing on top of the Suborder Anthropoidea, Infraorder Catarrhini evolutionary pyramid.

The Minimum span trees obtained from our analysis of the ventral plane samples is identical to the established evolutionary hierarchy illustrating the accuracy of this investigation. However, we could not correlate the same tandem with the minimum span trees of other domains. A number of factors can be responsible for this versatile character, but further study is necessary to shed more light on the topic.

The result obtained through this rudimentary investigation is quite appealing, but further revisions in this discipline is a vital requisite.

Index Terms: Form space, Asymmetry, Procrustes, Thin-plate spline, Semi landmarks, Cranium, Primates, Baboon, Chimpanzee, Common Marmoset, Gibbon, Gorilla, Human, Lesser Bushbaby, Orangutan, Ruffed Lemur, Squirrel monkey, Tarsier, Pseudo 3D, Principle components, Relative warp, Minimal Spanning tree.

1. Introduction

The phrasing of Morphometry had begun with the investigation of fossils of various individuals, including humans. The exact introductory thought itself was to assess the similitudes around different body parts for the evolutionary closeness around species.

This system has been successfully, however sparingly utilized for a long time to find out the evolutionary class structure around various taxonomic gathering of different period, various local areas and around adaptable populace collections. Different examination are embraced to center the effect of fluctuated ecological elements on this sort of variability and likewise the ensuing course of advancement.

Different examination groups have attempted to utilize this system to center the effect of fluctuated natural variables on the human populace, and this has led to the presentation of this method inside the therapeutic field for demonstrating the distorted pathologic body segments in individual. Ailments like schizophrenic psychosis, dementia and scattered numerous sclerosis causes the structures of the skull to modify. Utilizing morphometric investigation, this change will be effectively monitored and diagnosed. Morphometrics is an amalgam of arranged subjects, obliging profound understanding of imaging, scientific operations and connected math and measurable dissection.

The majority of the works as of nowadays is focused at a populace of a single animal categories or a solitary variety. Absence of extensive work including the discerning examination between various, evolutionarily related families stir our enthusiasm to undertake an analysis on them. Inside the extent of this dissertation, we wanted to analyze the Homo sapien skull of present existence in conjunction with numerous distinctive primates such a mandrill, Gorilla, Chimpanzees, Common marmoset, Gibbon, orangutan and attempted to lay emphasis on the association between them as indicated by their structure. Various progressed programming projects have been used in this methodology. We utilized the tps programming bundle suit particularly, tpsutil and tpsdig, embedding of the data points and for stacking of the pictures one-to-one. For factual investigation, Procrustes technique, Generalizes Procrustes Alignment, Principle segment Analysis, Thin-Plate Spline and various warps are acquired through PAST programming framework. Geometric Morphometric analysis typically employs these standard procedures.

The specimen samples could be successfully grouped in accordance with the shape familiarity among the skulls. Analysis in our flagship domain, along the ventral plane illustrated that the samples *0 (Baboon)*, 1 (*Chimpanzee*), 4 (*Gorilla*) and 7 (*Orangutan*) are morphometrically grouped together, while 6 (*Lesser Bushbaby*), 8 (*Ruffed Lemur*) and 10 (*Tarsier*) form a group of closely related species and the "rodent" like primates of 2 (*Common Marmoset*) and 9 (*Squirrel Monkey*) are in close proximity to each other. Homo *sapiens* sample is different from all other species due to variation in characteristics, and standing on top of the Suborder Anthropoidea, Infraorder Catarrhini evolutionary pyramid.

We did obtain some relatively optimistic results, as well were able to reveal and stumble upon certain unanswered queries which requires extensive further studies in this discipline that are based on versatile comparison and in diversified test environments. Due to the immense diversity of fauna along the world tropical belt, this will be the ideal location for similar morphometrics exploration.

2. Literature Review

Morphometrics, the estimation (metron) of shape (morph), is a subfield of facts with a history backtracking to the exact beginnings of this order. For instance, in 1888 Frances Galton presented the association coefficient and connected it to a mixed bag of morphological estimations on people. In 1907 he developed a technique to quantify facial shape that has later been termed two-point shape directions or Bookstein-shape directions. The provision of multivariate measurable methods, which were fundamentally designed during a major portion of the twentieth century, prompted alleged multivariate morphometrics. Anyways, in the 1980s, morphometrics encountered a significant insurgency through the creation of direction based strategies, the revelation of the measurable hypothesis of shape, and the computational acknowledgment of twisting lattices (for verifiable audits; Bookstein 1998; Rohlf and Marcus 1993; O'higgins 2000; Adams et al. 2004; Slice 2005).

The omnipresent provision of quick Pcs has introduced another dimension of information examination, allowing the investigation and visualization of expansive high-dimensional information sets alongside precise measurable tests dependent upon resampling methods. This new morphometric methodology has been termed geometric morphometrics as it jam the geometry of the landmark setups all around the investigation and accordingly allows to speak to factual approaches about real shapes or structures. About a few geometric methodologies to morphometrics, the Procrustes technique is the broadest and best seen in its numerical and measurable properties (Bookstein 1996; Small 1996; Dryden and Mardia 1998).

Other utilized morphometric routines are Euclidian separation network examination (Lele and Richtsmeier 1991, 2001), elliptic Fourier dissection (Lestrel 1982), and non-mark based methodologies like voxel-based morphometry (e.g., Ashburner and

Friston 2000), which is fundamentally connected in observance imaging. Various other universal morphometric methodologies are discussed through Blackith and Reyment (1971), Marcus (1990), and Oxnard (1983), or the assortment of strategies connected in histology (Baak and Oort 1983) and stereology (Weibel 1979; Baddeley and Vedel Jensen 2004). A thorough measurable correlation of a few of these strategies are presented by Rohlf (2000a, 2000b, 2003).

2.1 Scientific background Geometric morphometrics is dependent upon landmark coordinates. Bookstein (1991, p. 2) characterizes landmarks as loci that have names ('bridge of the nose', 'tip of the chin') and Cartesian coordinates. The names are expected to intimate correspondence (biotic homology) around structures. That is, landmark focuses have their areas as well as have the ""same"" areas in every other type of the example and in the normal of every last one of structures. These direction information can originate from a limitless exhibit of sources and can either be two- or three-dimensional. Two-dimensional directions are typically held utilizing a digitizing tablet or by measuring a picture on the workstation. Three-dimensional information might be caught specifically utilizing a direction digitizer, for example, a Microscribe or Polhemus, or may be measured on surface outputs or volumetric sweeps. Volumetric information are dependent upon picture cuts from Computerized Tomography (CT) or Magnetic Resonance Imaging (MRI) scanners—or their high-resolution variants µCT and µMRI (Sensen and Hallgrimsson 2009). These slices hold light black values that compare to tissue densities and are linked to acquire a three-dimensional representation of an item. Surface scanners give high-resolution 3D representations of an item's surface utilizing either laser or more universal optical innovation and may likewise incorporate composition data. Surfaces can additionally be concentrated from CT or MRI information. Most programming bundles permit landmark directions to be measured specifically on these virtual surfaces or volumetric items. Universal (i.e., non-geometric) morphometric methodologies ordinarily apply measurable procedures to an extensive variety of estimations, for example, distance and distance ratios, points, regions, and volumes. It is obviously conceivable to process the accepted inter landmark separations and plot from the landmark coordinates, however the first geometric relationship around the focuses may not be re-constructible from a specimen of choice, separation estimations. The crude landmark coordinates, on the other hand, can't be liable to measurable examination before dividing shape data from general size and "disturbance parameters" like position and introduction of the examples in the digitizing space that are not applicable for the investigation.

2.2 <u>Basics of Geometric Morphometric analysis</u> The principle objective of morphometrics is to study how shapes differ and their covariance with different variables. Despite the fact that morphometrics might be utilized to depict the type of any article it is for the most part utilized within science to portray creatures. Morphometrics is exceptionally imperative in science in light of the fact that it permits quantitative portrayals of living beings. Quantitative methodology permitted researchers to think about states of different living beings much better and they probably won't needed to depend on word depictions that normally had the issue of being deciphered contrarily by every researcher. This movement to quantitative portrayals was created by developments in measurable examination routines that permitted to translate gathered information. In the first place strategy for morphometrics called “traditional” morphometrics was carried out by measuring direct separations, (for example, length, width, and tallness) and multivariate factual devices were utilized to depict examples of shape variety inside and around gatherings. This methodology likewise some of the time utilized checks, proportions, regions and points measures. The greatest focal point of this system was that it was exceptionally straightforward, not-with-standing, it had a few difficulties. The most serious issue was that direct separation estimations are generally profoundly related with size and this makes shape examination grim. An alternate issue was that estimations taken from two dissimilar shapes could generate equivalent outcomes on the grounds that the information did not incorporate the area of where the estimations were taken with respect to one another. Furthermore, it was additionally not conceivable to recreate graphical representation of the shape from available estimations. Figure 1 delineates the issues of "traditional" morphometrics.

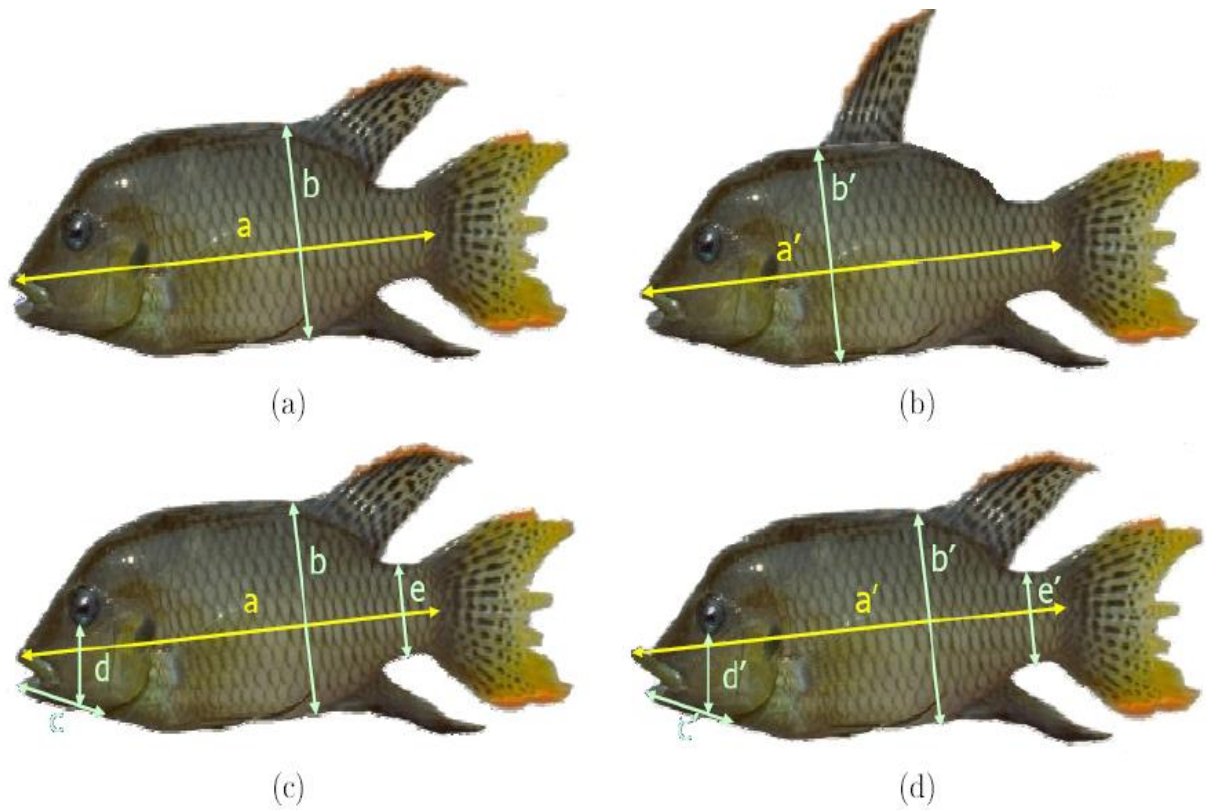


Figure 1: Size variables can be insufficient model of shape as seen in (a) and (b) a = a', b = b' that would lead to conclusion that shapes are equal. Length ratios can be used as seen in (c) and (b) the problem than is that a ≠ a' and lengths are compared b/a ≠ b'/a', c/a ≠ c'/a', d/a ≠ d'/a', e/a ≠ e'/a' conclusion shapes are completely different even though b = b', c = c', d = d', e = e'. Data from D. Adriaens, 2005.

To conquer these issues a more advanced system called Geometric morphometrics was made. Landmark-based geometric morphometrics utilizes a set of landmarks to depict shape. Landmark is a two or three-dimensional point depicted by a firmly defined set of tenets. The outcomes that are obtained by this strategy specifically rely on upon the nature of landmarks. A ton of efforts must be put to pick landmarks that might have high evolutionary significance. Every landmark likewise must be available on every examined life form. In the event that a landmark is not exhibit on no less than one of contemplated creatures it either must be stamped give or take or it cannot be utilized whatsoever. The amount of landmarks chose ought not surpass the amount of example specimens in light of the fact that additional landmarks will be excess. Typically number of landmarks is more or less equivalent to the amount of example specimens. There are three sorts of landmarks that could be utilized. Accurate landmarks that have some organic significance. Pseudo-landmarks are defined by relative areas e.g. "the purpose of most noteworthy curve of this bone". Semi-landmarks are demarcated by an area in respect to different landmarks e.g. "halfway between landmarks X and Y". Landmarks can at times have weighted esteem in investigation as stated by their significance.

2.3 Mathematical Basis of GM analysis Landmark information has a great deal of diversities in position, introduction and scale between examples. These non-shape varieties must be evacuated before further dissection. There are a few strategies used to superimpose landmarks each of them having different advancement criteria. The most straightforward one is two-point enrollment. This system deciphers, scales and turns all landmarks such that two named landmarks are in the same place in all examples. The greatest drawback of this system is that it evacuates all the information from those two landmarks. An alternate prominent system is Generalized Procrustes investigation (GPA, likewise frequently called generalized least squares). This strategy primarily figures the centroid of landmark conformations and makes an interpretation of it to the cause. This is carried out by taking k focuses in two dimensional space:

$$\left(\left(x_1, y_1\right), \left(x_2, y_2\right), \ldots, \left(x_k, y_k\right)\right)$$

The arithmetic mean of these points

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_k}{k}, \bar{y} = \frac{y_1 + y_2 + \ldots + y_k}{k}$$

And all the points are translated to the origin

$$(x, y) \rightarrow (x - \bar{x}, y - \bar{y})$$

Hence we obtain the following points

$$\left(\left(x_1 - \bar{x}, y_1 - \bar{y}\right), \left(x_2 - \bar{x}, y_2 - \bar{y}\right), \ldots, \left(x_k - \bar{x}, y_k - \bar{y}\right)\right)$$

Points are than scaled to a unit size. First centroid size is found

$$s = \sqrt{\left(x_1 - \bar{x}\right)^2 + \left(y_1 - \bar{y}\right)^2 + \ldots + \left(x_k - \bar{x}\right)^2 + \left(y_k - \bar{y}\right)^2}$$

And all the points are scaled.

$$\left(\left(\left(x_1 - \bar{x}\right)/s, \left(y_1 - \bar{y}\right)/s\right), \ldots, \left(\left(x_k - \bar{x}\right)/s, \left(y_k - \bar{y}\right)/s\right)\right)$$

Finally, the rotation is calculated by minimizing the sum of squared distance between corresponding landmarks. This step is complex. Two objects with point coordinates are obtained.

$$\left(\left(\left(x_1, y_1\right), \ldots, \left(x_k, y_k\right)\right), \left(\left(w_1, z_1\right), \ldots, \left(w_k, z_k\right)\right)\right)$$

One object is kept stationary and the other one is rotated around the origin so that the squared distance between points is abated. Rotation by θ angle gives the coordinates.

$$(u, v) = (cos\theta w - sin\theta z, sin\theta w + cos\theta z)$$

Distance between all points is calculated

$$d = \sqrt{\left(u_1 - x_1\right)^2 + \left(v_1 - y_1\right)^2 + \ldots + \left(u_k - x_k\right)^2 + \left(v_k - y_k\right)^2}$$

The need is to minimize this separation utilizing least squares technique. To continue it is important to determine the angle which provides the base squared separation. This requires to take the subordinate of d^2 regarding and θ tackling for θ when subsidiary is equivalent to zero. This gives

$$\theta = tan^{-1}\left(\frac{\sum_{i=1}^{k}\left(w_i v_i - u_i z_i\right)}{\sum_{i=1}^{k}\left(u_i w_i + v_i z_i\right)}\right)$$

All points of second object can then be rotated by θ angles with respect to the origin.

Following performing GPA to all specimens, shape differences might be concentrated by ascertaining differences in directions of relating landmarks. This information is than utilized as a part of multivariate examination to analyze shape varieties. Essential part dissection i.e principal component analysis, Canonical variates examination (CVA) and element investigation are some of usually utilized instruments. An elective system, i.e thin plate spline is utilized that permits to guide the disfigurement fit as a fiddle starting with one article then onto the next. This technique ascertains the alteration network that shows how one item could be twisted into an alternate. Typically objection is contrasted with the mean shape. Parameters of these malformation can then be utilized to measurably analyze varieties fit as a fiddle between populaces.

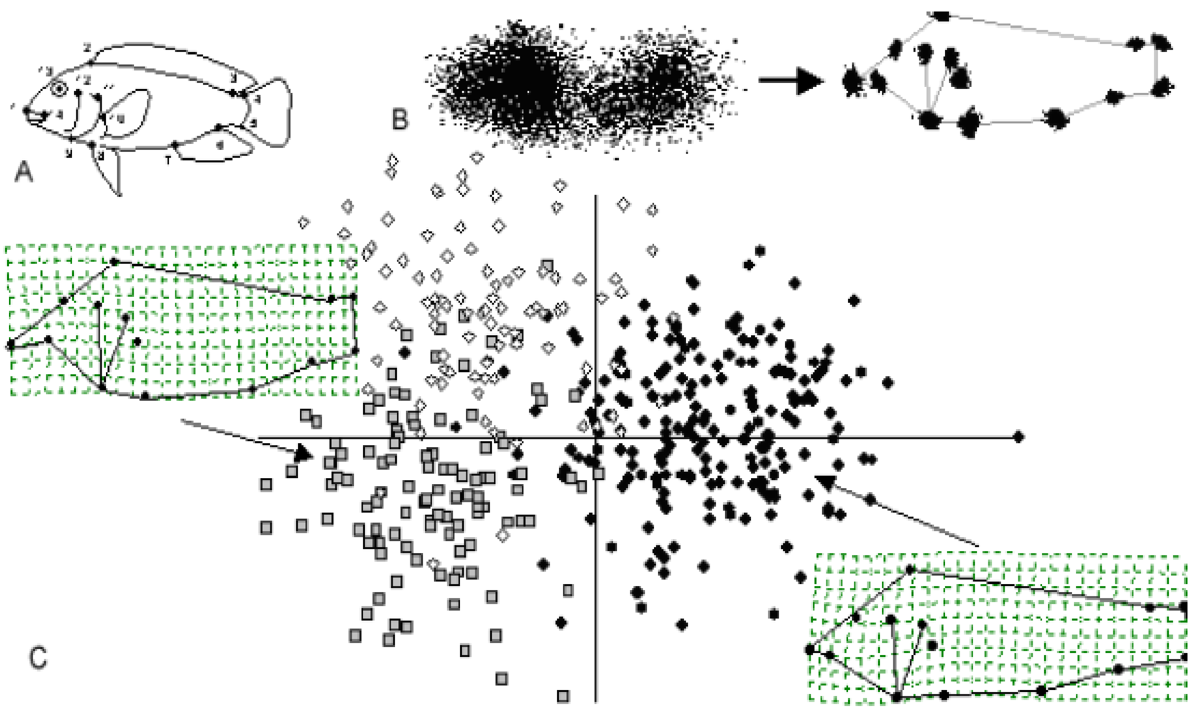


Figure 2: Graphical representation of the four-stage morphometric convention. A: Quantify crude information, B: Remove non-shape variety (data points of 412 specimens previously, then after the fact GPA), C: measurable examination (CVA) and graphical presentation of outcomes. Distortion frameworks for the mean example (amplified by 3x to accentuate shape contrasts).Data from L. Rüber and D. C. Adams, 2001.

The greatest focal point of geometric morphometrics is that it catches geometry of investigated protests, and conserve this data all around the investigation. This permits to see effects envisioned as measurable disseminate plots as well as conformations of landmarks focuses. Figure 2 show a sample of geometric morphometrics dissection.

2.4 <u>Missing Data in Morphometrics</u> From a factual perspective, landmarks that can't be measured in light of the fact that anatomical structures are severed or disfigured are termed "missing information". As geometric morphometric techniques oblige all measured examples to have the same number of landmarks at homologous positions, the variables or the cases with missing qualities must be either prohibited from the examination, or the missing qualities need to be evaluated dependent upon the accessible information. There exists a huge collection of writing on strategies for evaluating populace parameters, for example, means or relapses in the vicinity of missing qualities (Schafer 1997; Mclachlan and Krishnan 1997; Schafer and Graham 2002). In morphometrics, especially when connected to the paleo-sciences, the evaluations of the missing qualities may be of investment themselves, e.g., as an approach to recreate divided fossils. In Gunz (2005) and Gunz et al. (2004, in survey) have laid out a schema for assessing missing shape directions, misusing the high repetition of shape variables—particularly when the estimation focuses are nearly separated and considering former natural learning about useful and geometric imperatives, symmetry, and morphological joining.

The most effortless yet most capable approach to manage missing and distorted anatomical structures is to restore two-sided symmetry. At the point when parts are absent on both sides or in the symmetry plane, one can foresee the directions of the missing landmarks utilizing data from complete cases. Measurable recreation is pointed at improving the factual probability of the assessment, utilizing an expansion of a regular missing information calculation, the desire amplification strategy (Dempster et al. 1977), for Procrustes shape variables. Geometric recreation utilizes the smoothness properties of the thin plate spline to gauge the missing directions dependent upon a solitary reference design. A TPS addition is registered dependent upon the subset of landmarks that are accessible in both the reference and the inadequate example. This interjection capacity is utilized to guide the missing landmarks from the reference onto the fragmented target, setting the landmark evaluates so that the misshaping between the reference and the deficient example is as smooth as could be expected under the circumstances (least twisting vitality). The reference design might be a real example or a Procrustes aggregation normal and may look like the fragmented examples in properties, for example, bunch alliance, age, or geographic starting point. Characteristically, picking diverse references will bring about marginally distinctive recreations. Whether comparable determinations may take after from an assortment of reasonable elective reproductions is testable in the process of the investigation: the appropriation of distinctively remade shapes reflects the lack of determination because of the missing information and the affectability to earlier presumption.

The greatest burden of landmark-based geometric morphometrics systems is that various landmarks accessible can some of the time be insufficient to catch the state of an item. An elective is to utilize framework dissection technique. This technique first separates a limit around an article. Focuses are than digitized along the limit. These focuses are than fitted with a numerical capacity (generally some manifestation of Fourier examination). Different bends are than looked at utilizing coefficients of the capacities as shape variables in multivariate dissection. This methodology however additionally has a few constraints. This strategy is not fit for catching shape changes inside an article (landmark-based system can have landmarks not just on the limit). This technique is likewise difficult to apply when investigated information is three dimensional.

Geometric morphometrics is an animated examination territory and there are new techniques proposed to address the limits of strategies presented previously. One especially guaranteeing technique means to make a strategy that might join landmark-based and framework dissection into one system taking focal points from both. This technique proposes the utilization of sliding semi-landmarks that are made between genuine landmarks yet on the limit of an article. This system can likewise be embraced to be utilized with three dimensional information. The most serious issue of these new systems is that there still is no generally affirmed procedure on the best way to utilize them.

2.5 <u>Use of Morphometrics in versatile studies</u> Morphometrics processes are functional to commence various biological studies based on evolutionary changes, to study the alteration between entities and/or variation in their segmented parts, impact of various environmental factors or geological elements on specific species or to analyze any injury. The 2D or 3D shape disparities of various entities have been recorded and statistically examined by the virtue of numerous. A software which can visualize and scrutinize the shape discrepancies in 3-Dimensions have been designed by comparison of human, chimpanzee and baboon morphologies. The taxonomic prominence of Homo erectus sensu lato had been challenged in around 1980s. It was projected that the early African fossils may signify a completely different species termed H. ergaster. 3D Geometric Morphometric analysis was extensively used to resolve this difference in opinion, and enumerate the global shape disparity in the cranial vault of the H.erectus. The outcomes of the geometric morphometric analysis illustrated that this Homo erectus has feature characteristics that are remarkably similar to a particular subspecies of popionin monkeys. Analysis by biological anthropologist Karen L. Baab, in the year 2007, indicated that there may be a trivial variation in the characteristics' governing features between the Homo erectus and the sensu stricto, if any. Various attempts have been undertaken to recreate the evolutionary pathway on the basis of variations in the cranial morphology.

Various facial features and cranial structures are modified by exposure to versatile environments. These alterations in turn leads the evolutionary process. The shape of the head can be digitized and the form can then be correlated with the climatic conditions from the indigenous locality of the specimen. The variation of the facial features in accordance with the change of climatic conditions have been broadly documented (Katrina Harvati et al., 2006).

Three dimensional GM analysis have been utilized to attempt a quantitative dissection of human mandibular shape. The primary object of this study was to record the environmental and geographical impact on the structure manifestation of the lower jaw of humans. A research team from Northwestern University's Feinberg School of Medicine is trying to expand this research and employ the GM analysis in its pursuit to quantitatively dissect the "Neanderthal" mandibular features. Significant geological patterning has been observed from the analysis of the late human mandibular shape , as well as a few parts of mandibular morphology reflects an atmospheric perspective and others utilitarian specialization (Elisabeth Nicholson et al., 2006).

This method of analysis has been proposed to be used to evaluate the effect of environmental conditions on the human population and how is manages to shape the present and future. The introduction of this methodology in the medical sciences and medical diagnostics is based on this proposal, and as of now it is used to compare, visualize and record various deformed body parts of Homo sapiens. Claims have being made that by utilization of this process, various Shape altering diseases can be detected, the symptoms observed, recorded and the patient diagnosed.

2.6 Landmark selection

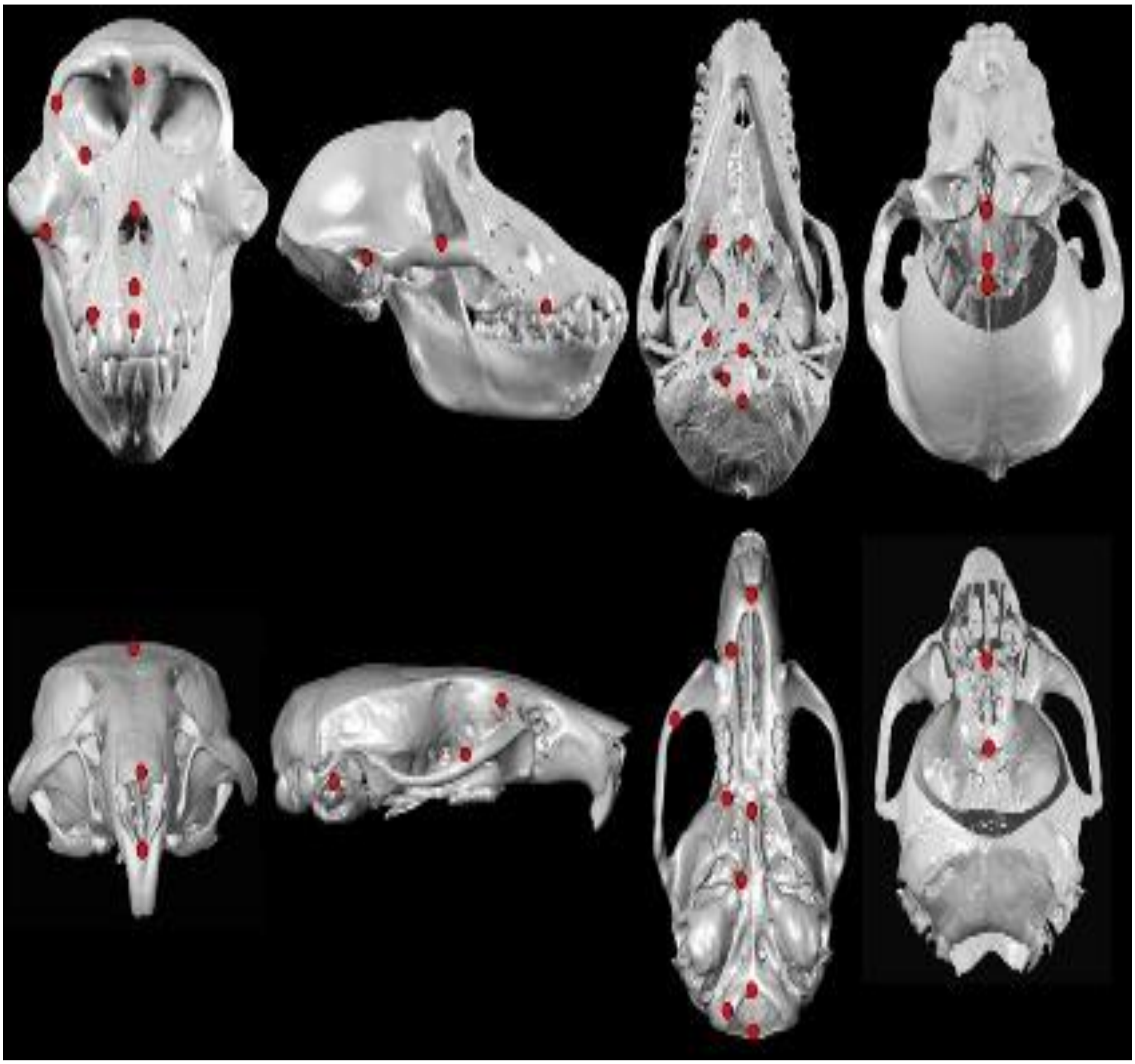


Figure 3: Three Dimensional Landmarks established on Cranium samples of two different species.

Landmarks are used to digitize a specific shape so that the sample can be statistically analyzed. This process is commenced by specifying a specific region of interest, bounded by the induced landmarks, containing coordinated data sets. For all intents and purpose, each morphometric examination is accomplished by using such data points. If the specific structure is simple, a few landmarks will be sufficient for proper analysis, however in case of a complex or a multivariate model, the efficiency of the process increases with the density of the landmarks placed. Ideally, higher the number of Landmarks introduced, the obtained results are more accurate. Determination of Points of interest, their size and their accumulation density on an example's examination. An investigation is significantly influenced by the detail of Landmarks, their arrangement and their fixation.

2.7 **Statistical Analysis** When the morph of the specimen is digitized by the landmarks embedded, fitting examination of these specimens are needed for viable correlation around the cases. There are different courses by which this could be attained, however the most well-known and the best of the methodology are universally utilized within morphometric examination. Some of these are:

2.7.1 **Procrustes superimposition** Mathematically denoted as the smallest value of the aggregate-squared detachment between any two data points. This value is obtained by taking into consideration all the characteristics inputs over all the size based accredits like rotation, scales and linear translations. The area of interest produced by this methodology, does not have any backing for immediate factual and numerical dissection, because of its nonlinear shape conformity (David F. Wiley et al., 2005, Karen L. Baab 2008, Katrina Harvati et al., 2006, Elisabeth Nicholson et al., 2006).

2.7.2 <u>General Procrustes Alignment (GPA)</u> A set of accord point of interest is chosen to lessen the aggregate squared distinction between the adjusted information data sets and the covenant conformity. This procedure beats the issue experienced through Procrustes analysis, and produces a straight space which backs immediate geometric investigation to be incorporated. We can additional, directly interject our set of info design, once the General Procrustes Alignment is accomplished (David F. Wiley et al., 2005, Karen L. Baab 2008, Katrina Harvati et al., 2006, Elisabeth Nicholson et al., 2006).

2.7.3 <u>Principal component Analysis (PCA)</u>

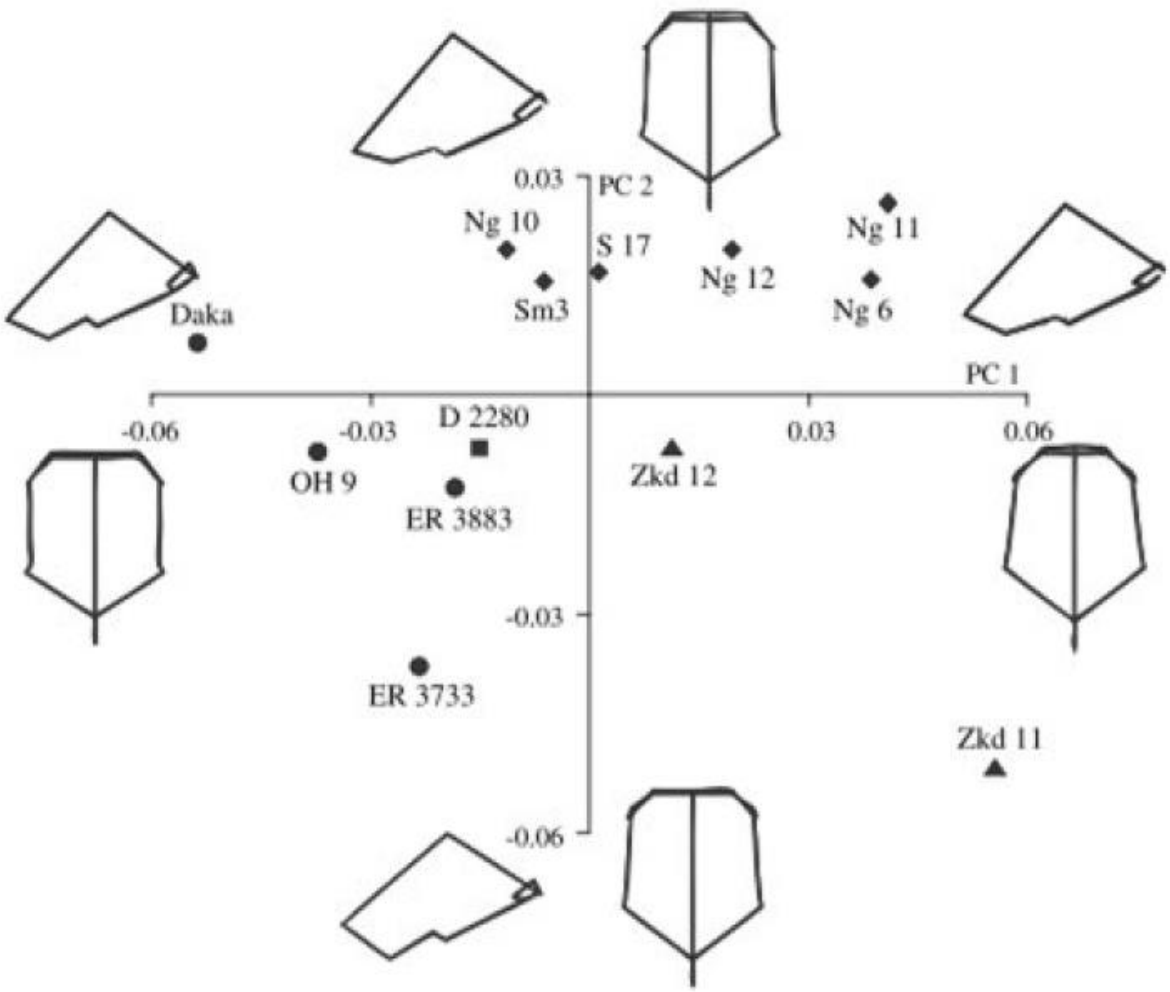


Figure 4: Principle Component analysis (PCA) plot between PC1 and PC2.

To break down the neurocranial shape diversity in this taxon, a Principal Component Analysis was implemented on the 16 and 32 data point sets for H.erectus specimen. The African and Asian fossils are isolated by the smaller landmark set along the PC1 which is 34.6% of the summative alteration.

2.7.4 <u>Thin-Plate Spline (TPS) and Warps</u> Exploration of form characteristics is undertaken through this process, an optional, but regularly utilized system. Structuring of a TPS warp requires an information set of landmarks and the set of objective oriented point of interest. Proficient characterization of TPS warp requires these prerequisite and it permits the data points to agree with the objective set. The source surface cross section is disfigured till the objective shape is accomplished.

3. Material and methods

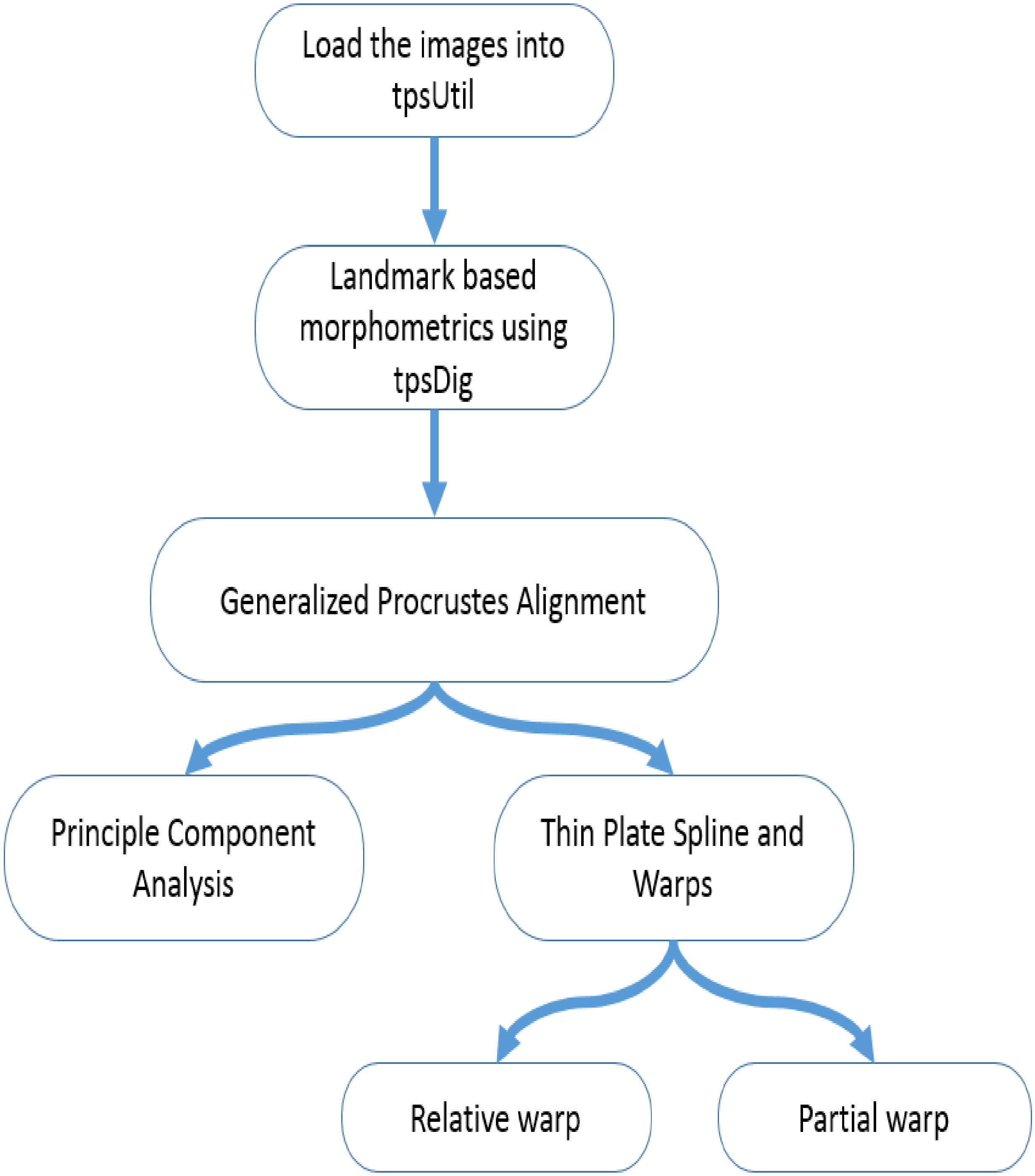


Figure 5: Flowchart of overall work, segmented into various important steps.

3.1 Important facts based of the Primates used as samples

The following section includes the basic facts about the specimens used in this investigation, in clockwise manner from top-left Sample profile, skeletal profile and Indigenous geographical locations.

BABOON:

Baboons live in African savanna Woodlands, and survive on tubers, grasses, roots, seeds, and fruit, but then again hunt small animals, generally smaller than themselves. The three classes of baboons, Theropithecus, Mandrillus and Papio, are not closely related and has variable morphology and behavioral characteristics. Papio cynocephalus has a characteristic mutt-like muzzle and a coat of fur illustrating a tannish shade of yellow. The yellow baboons are sexually dimorphic, weighing approximately between 11-26 kg (24-57 lbs), with females approximately 1/2 the size of males. Generally, baboons live in polygamous social clusters with up to nine females, but they occasionally also flock into foraging "bands" of 30 individuals, or into "troops" of as large as 100 members.

*

CHIMPANZEE:

Chimpanzees are indigenous to the African Locations. They are generally omnivorous as they feed on both shrubbery and meat. *Pan versus*, *P. schweinfurthii*, *P. paniscus*, and *P. troglodytes* are the four extensively documented subspecies of chimpanzees. *P. troglodytes,* generally characterized by large ears, have a prognostic (protruding) face with large lips and pronounced supraorbital crests. Males typically weigh 34-70 kg (75-154 lbs). Females weigh up to 25-50 kg (55-110 lbs) with higher weights usually observed in captivity. Chimpanzees are also known to manufacture and use tools, something that is frequently considered as an exclusive human trait.

MARMOSET:

Marmosets, live in the New Worlds, the tropical forests of Southern America. The common marmoset, *Callithrix jacchus*, weighs an average 370 gm (.81 lb.). *C. jacchus,* gray colored, has a banded tail. It has distinct, long tufts of white/ cream fur around the ears. Characteristically, these are small, fast, hostile, territorial, and depend on a variety of insects, fruits, and vegetable for food. The marmosets have modified lower canine teeth, which are improvised to penetrate tree barks, enhancing their ability to access the tree exudate (tree gum).

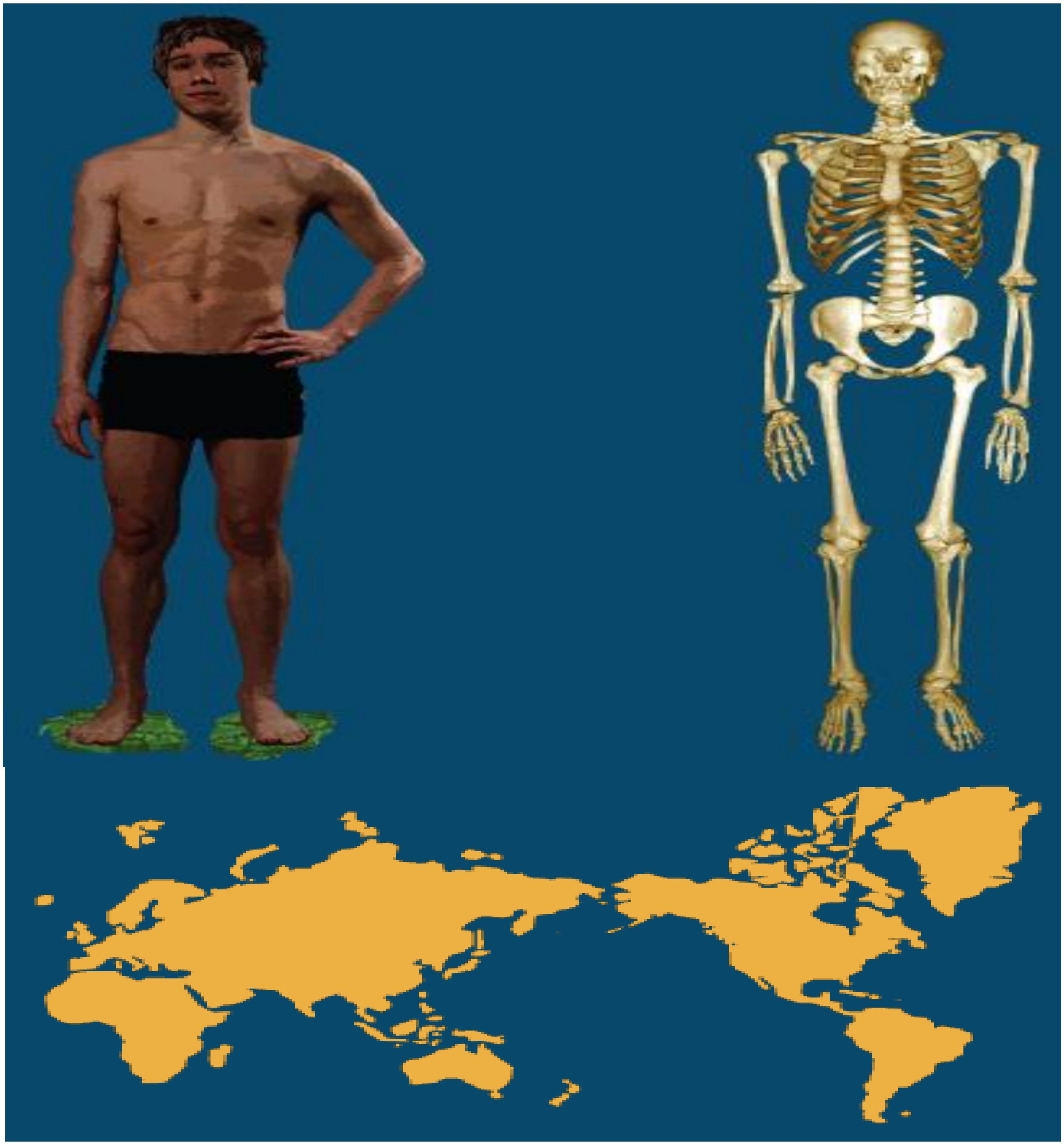
GIBBONS:

Gibbons, are an energetic species. This superfamily includes 10 subspecies in the family-Hylobatidae. Generally these Measure about 75-90 cm (29.5-35.4 in) in length and weigh 8-13 kg (17.6-28.6 lbs), Siamang (*Symphalangus syndactylus*) has a long arm-spread length of more than 1.5 m (almost 5 ft.). *S. syndactylus* has a black coat which is longer and thinner than those of the other species. Visually recognizable by their large, hairless, empty air sac in front of their throats. This sac activates and swells during vocalization. Gibbons have a webbing between their second and third toes, are brachiators, and travel through the trees by swinging with their arms for leverage.

GORILLA:

Gorillas are the largest living primates, reach up to a height of 1.75 meters/ 70 in and walk on two feet. Adult male gorillas weigh about 135-275 kg (297-605 lb.) and females, 70-140 kg (154-308 lb.). They have an arm span length of up to 2.75 meters. In captivity, gorillas can weigh a lot more with statistics indicating values as high as 350 kg (770 lb.). The three subspecies within the genus, *Gorilla gorilla* indigenous to the western lowland; *G. gorilla graueri* aboriginal to the eastern coastal area; and *G. gorilla beringei* commonly the mountain gorilla. Primarily vegetarians, Gorillas are known to easily consume tough plant material.

HUMANS:

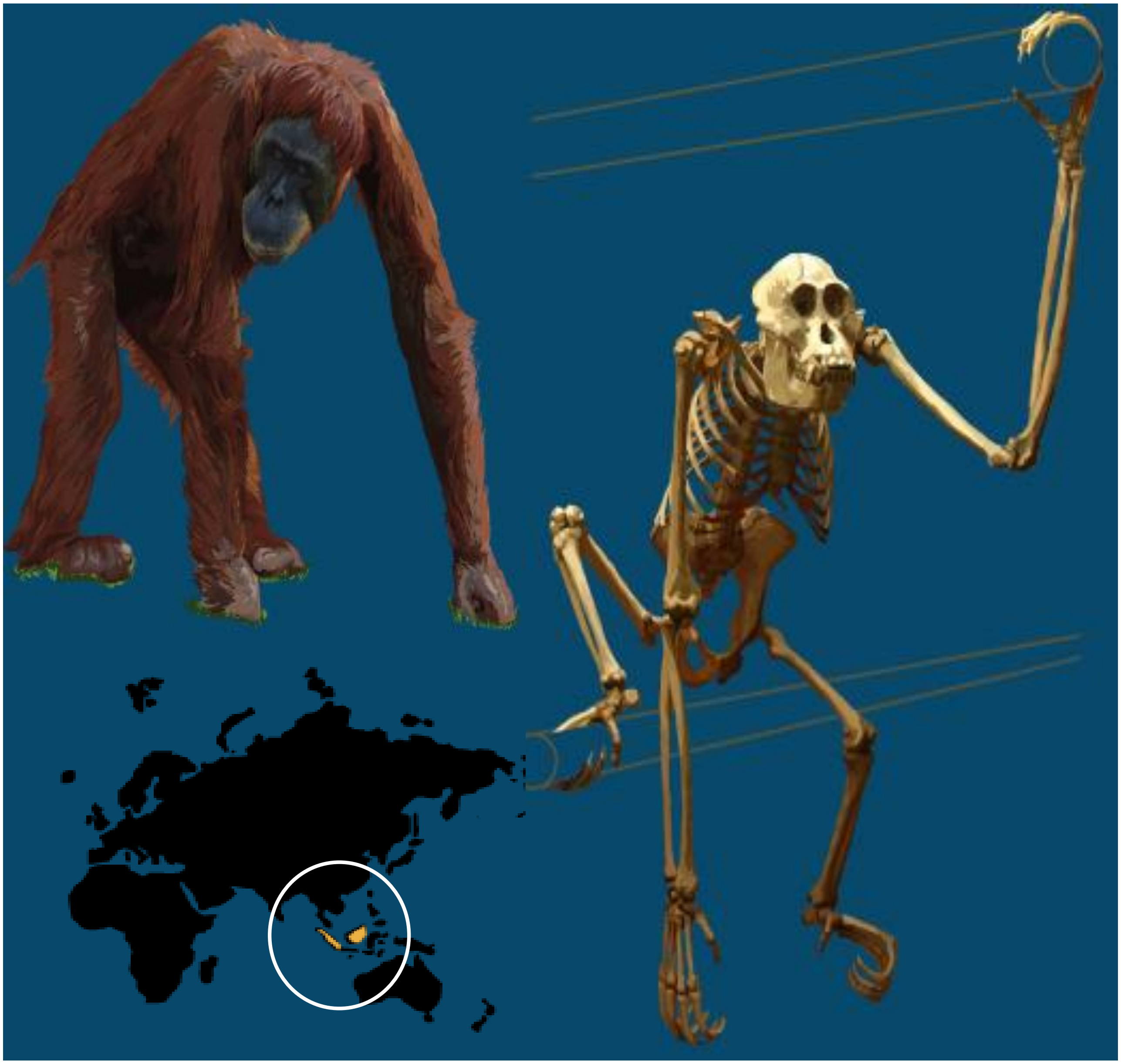


Humans are the only primates that practice obligatory bipedalism and their skeletons show characteristic modification for this type of movement. Humans' can regulate their body temperature and possess a large brain that is technologically advanced. This allows for expertise, and multiplicity of culture and language. Humans live permanently in almost all habitable terrestrial locations of the planet, and temporarily occupy inhospitable environments such as the Sahara and the Antarctic.

LESSER BUSHBABY:

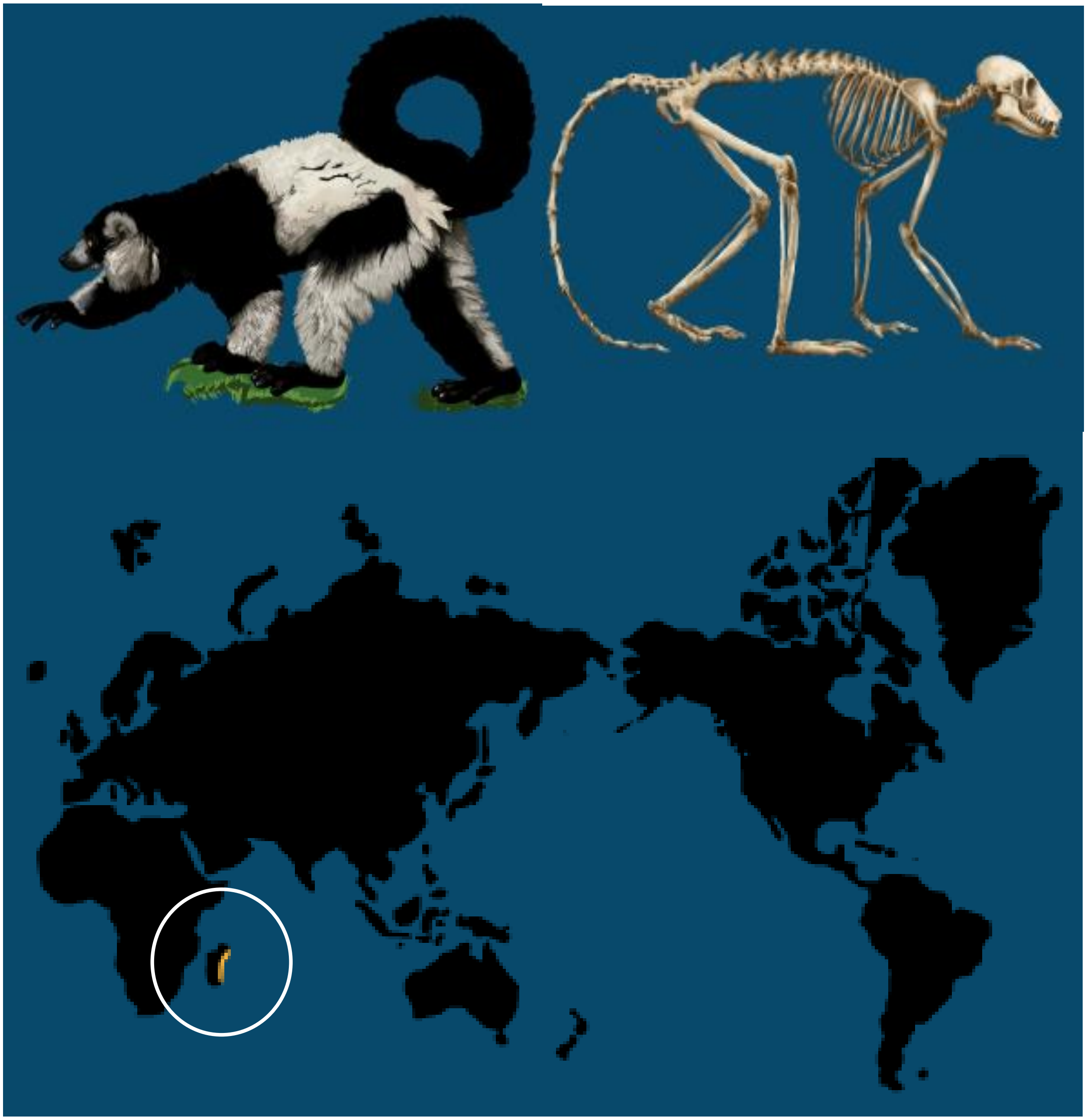
Gallegos are African arboreal, nocturnal primates. There are three species - *Galago senegalensis*, the Lesser Bushbaby; *G. crassicaudatus*, the Thick-Tailed Galago; and *G. Alleni*, Allen's Galago - living in a variety of habitats. *G. senegalensis* is commonly found in the coast of Guinea, the savannas of West and seldom in sub-Saharan Africa. Relatively, a smaller species, Males weigh between 151-173 g and females 150-163 g. *G. senegalensis* has a characteristics stripe between the eyes extending down the nose, and ears with four transverse fold lines that the animal uses to compress or outspread its ears. Gallegos are possessed with longer forelimbs, compared to their hind limbs. This characteristic feature enables them to cross distances of more than a couple of meters in a single jump.

ORANGUTAN:



Orangutans have a place with a solitary animal types recognized as Pongo pygmaeus arranged into two strains: P. p. pygmaeus found in Borneo, and P. p. abelii existent in Sumatra. They have committed upper appendages adjusted for suspensory movement, yet are a tetrapod when on the ground. Reaching to heights about 1.25-1.5 m (4 - 5 ft.), Orangutans have a normal arm compass length of something like 2.25 m (7.4 ft.). Adult males of the clan weigh something like 50-90 kg (110 - 198 lbs) while grown-up females more or less measure about 30-50 kg (66 - 110 lbs). They have jutting noses, a protruding temples, dainty lips, and little ears. Orangutans have coarse, slender layers that are earthy-red in color, however P. p. abelii has a marginally lighter coloring. Naturally frugivores, Orangutans every so often devour bark, leaves, and fowl eggs.

RUFFED LEMUR:



Lemurs, belong to the family of *Lemuridae*, comprising of nine genera and three intermittent subfamilies, including the genus *Varecia* and the species Â*V. Variegata*, the Ruffed Lemur. The Black Lemur also known as *V. v. variegata* give the impression of a black coat, with intermittent patches of white fur, while the Red Lemur (*V. v. ruber*) has a sporadic coat with areas of fur, black in color. *V. variegata* is the largest species in the family weighing about 3.2-4.5 kg (7-9.9 lbs) and having a torso length, inclusive of the tail as 56-65 cm. *Varecia are* inhabitants of the forest of tropical Madagascar, which are home to primarily all Lemurs. *V. variegata* is principally frugivorous and is crepuscular, implying that it is generally more active during dawn and dusk.

SQUIRREL MONKEY:

Squirrel Monkeys are indigenous to the forests of Central Latin America. Belonging to *Cebidae,* the family consist of 65 species spread over 11 genera. These individuals are well-known for their mask-like white face topped with a dark cap, and a hairless muzzle around the nose and mouth. They have gray bodies, a white belly, and have a dual-toned tail with strips of black. *Saimiri*, the smallest of the geneses, weighs around 500-1100 g (1.1-2.4 lbs) and has a top-to-tail characteristics span of 35-42.5 cm (13.7-16.7 in). *Saimiri oerstedii* can be found on the west shoreline of Costa Rica and Panama surviving on flowers, fruits and insects. *Saimiri* are characteristically termed as "flush-hunters", signifying that they catch their prey that travel, when other monkeys migrate through the forest.

TARSIER:

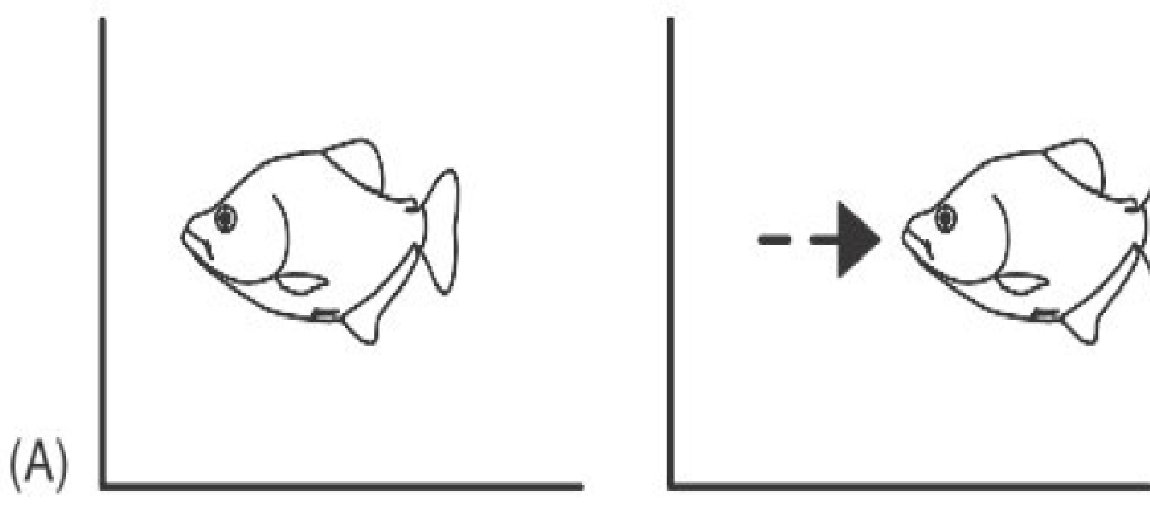
Tarsiers' have stationary eyes requiring them to rotate their heads through 360 degrees to change their arena of vision. Of all the related species, Tarsiers have the biggest eyes to body ratio with a typical eyeball diameter ranging up to 16 mm. *Tarsius bancanus*, one of the most important *Tarsius* species, is found in Borneo and Sumatra. Adults weigh roughly 80-165 g (.17-.36 lb.), and have head to body extent of 8.5-16 cm (3.3-6.3 in). They have relatively small upper bodies, extremely long hind limbs, and relatively massive heads. Tarsiers are nocturnal, feeding primarily on invertebrates and vertebrates, including small poisonous snakes. Anatomically augmented to act as Vertical clingers, jumpers and leapers, they can cover distances of up to 6 m (~18 ft.) at any time.

Courtesy of the information: (a) http://www.eskeletons.org/ (b) http://www.efossils.org/

3.2 Important Terms Preceding a detailed portrayal of the materials and methods exploited for this investigation, a basic comprehension of expressions are a requisite:

3.2.1 Size Size is on the whole characterized as the extent of the degree of a living being, measured long along any entity axis, area, region, volume and weight. It can additionally be designated as a direct blend of all the calculable amounts which might be cohorted and/ or corresponded in some structure or the other. Changes in size emerges with a slight variety in the scale, pivot and the area in the picture included as the region of interest. Morphometrics Investigation is focused around the idea of shape keeping it completely autonomous from disparity in size impact components, for example, scale, rotation, projection, translation, turn and interpretation.

3.2.2 Shape In light of morphometric analysis, form or 'shape' could be characterized as all the linear measurable entities including the ones gained in the wake of separating out area, scale and rotational impacts from the object of investment.



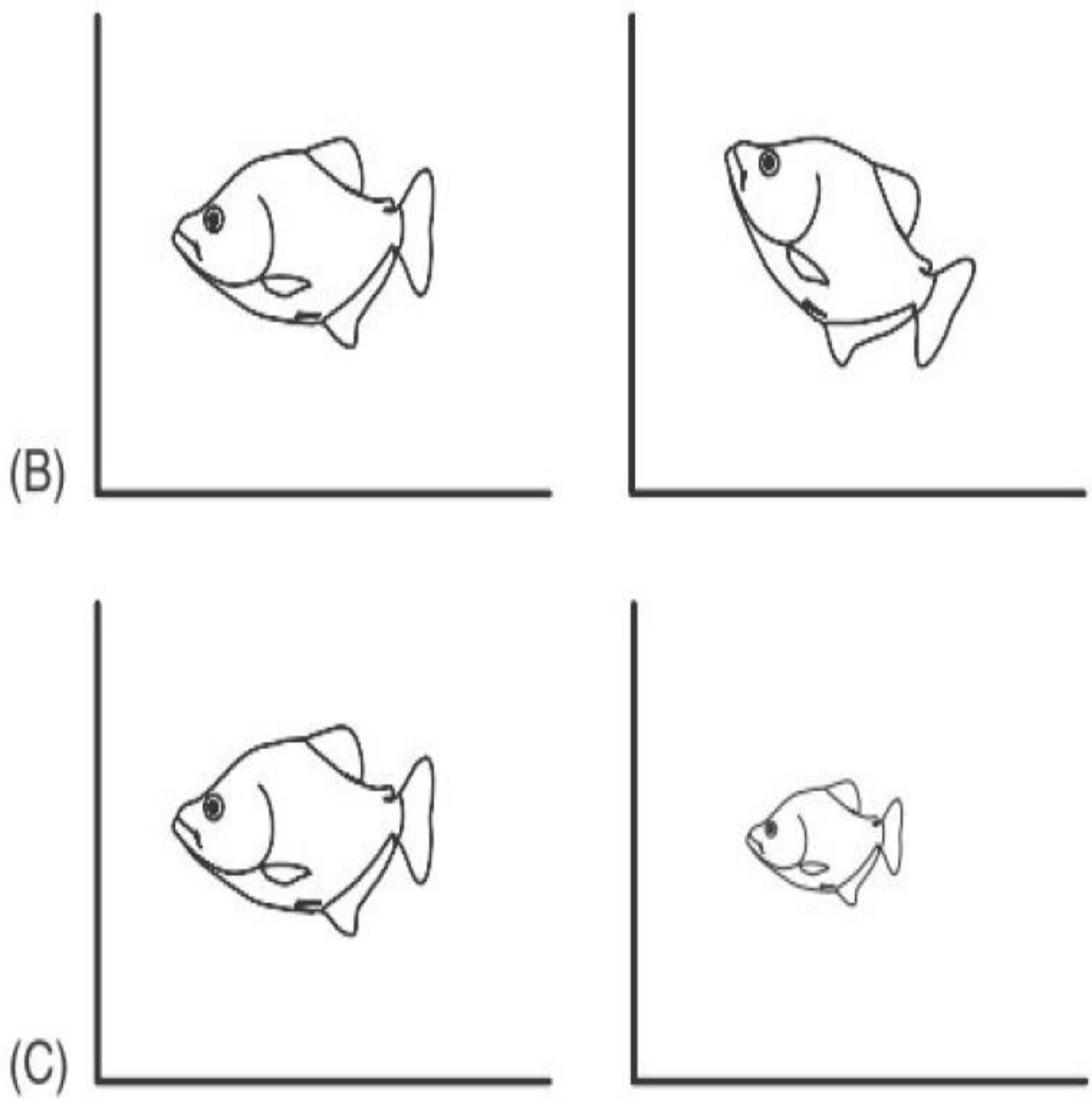


Figure 6: Size effect of (A) location, (B) rotation and (C) scale are illustrated.

3.2.3 **Landmarks** Landmarks or structural Milestones could be characterized as a set or sets of homologous anatomical loci that doesn't alter their structural areas or locations in respect to the other insertion points. These need to give adequate scope of the morphology and might need to be discovered and induced repetitively, consistently and reliably. For the most part, the Landmarks are placed such that they lie on the coherent plane.

A discrete set of K correspondent data points are injected on ‘N’ input article surfaces;

$$L^i = (x1, x2, x3\ldots.xK),\ 1 \leq t \leq N \qquad (1)$$

Greater the number of points, reporting of the sample features will be improved and an effective sampling of the shape can be obtained.

3.2.4 **Mean** For an information set, the terms mean, mathematical expectation, numerical expectancy, and average are employed synonymous to allude to a focal worth of a discrete set of numbers: particularly, the aggregate of the qualities divided by the total amount of qualities, or as it defined “the sum of items divided by the number of items”;

$$\ddot{X} = \sum_{i=1}^{n} Xi/n \qquad (2)$$

Where $\ddot{X}$ is mean, X is the magnitude of item and n is the quantity of items.

3.2.5 **Standard Deviation** In statistics and probability premise, the standard deviation (denoted by the Greek letter sigma, σ) demonstrates the amount variety or scattering from the normal entity. A low standard deviation shows that the information obtained from the data sets tend be near the mean (additionally called expected quality); a high value of σ demonstrates that the information focuses are spread out over an expansive extent of qualities.

$$\mathbf{S} = \sqrt{\sum_{i=1}^{n} \frac{(Xi - \ddot{X})^2}{n-1}} \tag{3}$$

3.2.6 **Variance** Variance measures the separation in terms of distance within a discreet set of data points. A difference of zero demonstrates that all the qualities are indistinguishable. It is always a non-negative entity, just in conjugation with the physical distance. When data points are close to the mean, their characteristics is represented by a value of variance. Being close to the mean implies the points are close to each other, illustrating an exceptional presision of the system, while a high variance indicates that the data points are very spread out from the mean and from each other. Variance is the square of standard deviation

$$\mathbf{S}^2 = \sum_{i=1}^{n} \frac{(Xi - \ddot{X})^2}{n-1} \tag{4}$$

3.2.7 **Covariance** covariance is a tool for measurement of the change in two random variables, when both entities undergo change simultaneously. If the variables tend to show similar behavior in terms of the magnitude of their respective values such that a greater value of one variable corresponds with the greater values of the other entity, and the same holds for the lesser values, i.e., the, the covariance is positive. This tends to illustrate that the components display similar properties.

$$\mathbf{Cov(X,Y)} = \sum_{i=1}^{n} \frac{(Xi - \ddot{X})(Yi - \ddot{Y})}{n-1} \tag{5}$$

3.2.8 Eigenvector and Eigenvalue: Eigenvectors are a special set of vectors associated with a linear system of equations (i.e., a matrix equation) that are sometimes also known as characteristic vectors, proper vectors, or latent vectors (Marcus and Minc 1988, p. 144). The non-zero vectors, A, satiating the equation :

$$(\mathbf{X} - \lambda \mathbf{l})\mathbf{A} = \mathbf{0} \qquad \textbf{(6)}$$

The eigenvalues of X, is provided by λ, satisfying the above equation. Eigenvectors are perpendicular to each other and provides the most diminutive set of axis for a vector space.

3.3 <u>Procrustes fitting (Superposition)</u> Procrustes distance is the smallest value of the aggregate-squared detachment between any two data points. This component is calculated over all rotations, translations and scales.

$$\mathbf{D}\,(\mathbf{L^i}, \mathbf{L^j}) = \sqrt{\sum_{n=1}^{K} (ln^i - \ Lm^j)^2} \qquad \textbf{(7)}$$

This type of fitting overlay all the data points point sets of every last one of pictures with the evacuation of measurement impacts, for example, rotation, pivot, scaling and translation. It structures a non-direct space in which it exacts a shape on the space of point of interest conformity. Then again, this sort of pair-wise superimposition can't transform a compound reciprocated common arrangement of all the point of interest sets, which suggests that, in the event that we have a tripicate of milestone specifically, La, Lb and Lc and we have effectively adjusted La to Lb and Lb to Lc, doesn't mean La will be attuned to Lc. Moreover, measurable investigation can't be associated to this sort of non-linear space. In this study, we have attempted the GPA methodology to reverse this disadvantage. In this incomparable system, a specific milestone set is decided to diminish the total squared qualification between this particular milestone set and the imported made data sets. This system changes the nonlinear space to an immediate space so that the measurable strategies can authentically associated.

3.4 <u>Principal Component Analysis (PCA)</u> This is a viable apparatus, used to center the adaptations in the data and to express the data to highlight the likenesses and contrasts between them. It's an instrument for getting the speculative variables called standard parts to consolidate the multivariate difference around the peoples. These factors are rectilinear combination of the starting variables. Principal Component Analysis basically diminishes the measure of estimations in the wake of uncovering the structure and in this way cinches the data without any earth shattering misfortune of measurable material.

Principle Component Analysis (PCA) is an accurate framework that uses orthogonal adjustment to change over a set of discernments of possibly related variables into a set of characteristics of straightly uncorrelated variables called Principle Component. The degree of Principal part is small to the measure of unique variables. This change is described in such a way, to the point that the first key part has the greatest possible contrast (that is, records for however a great part of the variability in the data as could be normal), and every one succeeding segment accordingly has the most extreme raised vacillation possible under the commitment that it is orthogonal to (i.e., uncorrelated with) the previous segments. Primary segments are guaranteed to be freed if the data set is intermittently passed on. PCA is sensitive to the relative scaling of the first variables.

The enclosed step illustration exemplifies the steps involved Principal Component Analysis:

Step 1: Data is arranged in a table format, such that the coordinates (x, y) are placed in the column and the objects in rows.

Step 2: The mean is subtracted from each of the dimensions.

Step 3: The covariance matrix is calculated.

$$C = \begin{bmatrix} cov(xx) & cov(xy) \\ cov(xy) & cov(yy) \end{bmatrix}$$

Step 4: The eigenvectors and the eigenvalues are computed for the covariance matrix

Step 5: The eigenvector having the highest eigenvalue is the principal component of the data set. The eigenvectors can be arranged in a descending order based on eigenvalue.

Step 6: The eigenvectors having the large eigenvalues are kept and the rest are discard. In conclusion these eigenvectors are arranged in a matrix form, named feature matrix.

Feature Matrix = (eig1 eig2......eigN)

Step 7: To conclude the transpose of the feature matrix is obtained and is multiplied to the transposed mean subtracted data (step 2). The data we obtained finally is the original data in terms of the vectors we selected, meaning we reduced the dimensions by discarding low eigenvalues.

3.5 <u>Thin Plate Spline and Warps</u> Thin-plate spline is an extension function used to focus the general contrast of form between a reference and the specimen shape. Spline and warps analyze shape conformities on the whole and not only of the points of interest. Streamlining the idea, one can say that the procedure of thin-plate spline is a modification of the source shape so as to distort it to the target shape by putting it on a tetragonal framework grid and disfiguring the matrix and consequently the shape is distorted until the object shape is realized. This change in conformation of the grid and the associated form adjustment provides a congenial awareness of the disparity in the contour profiles.

3.5.1 **Principal Warps** A multidimensional space might be organized from a specimen profile, where all the conceivable rotations of the model could be plotted. This space is usually called deformation or distortion plane. These bearings not just help to improve the estimation of movement of points from the source, however also describe consummately the orthogonal fixed groups of spacial linear functions which is grounded on the "twisting vitality" of the alteration. Such premise capacities are Principle warps. Subsequently, Principle warps may be described as a domain of all conceivable changes a long way from the basis shape and are not destitute of any objective similarity. Principle warps is utilized to focus the partial warp.

3.5.2 **Partial Warps** Separated from primary warps that are simply subject to the source setup, partial warps are explicit compared to a source, and in like manner for a specified target. In a gathering of samples, if the average is obtained as the source, then every model will have particular explicit partial warp depending on the spin from the moderate shape to the particular sample. A change is totally and especially portrayed as the biased entire of the partial warps and the relative portion which is the zero order warp and depicts the square distortions of scaling, cutting, developing and shearing. The central normal for partial warps is that, are organized according to their close-by proportionate distortions. The nearby bends are related to the higher order warps, while the large scale and global contortions are embraced by the primary mandate warp.

3.5.3 <u>Relative Warps</u> The principal units of a collection of thin-plate spline alterations are the relative warps. Relative warps are their self-specific modifications and might be envisaged by the matrix distortion. They might be utilized as an instrument for information diminishment, for a set of conversions by putting the first changes as focuses in the scatter plot.

3.6 <u>Softwares Used</u>

3.6.1 <u>TpsUtil</u> This is a component of the Tps Utility suite is used to build a tps file before the process of collection of data is commenced. The Tps file is built by compiling a list of specific required specimens.

Methodology for the use of tpsUtil are elaborated in the underlying steps:

Step 1: All the specimens/images are placed in a single folder.

Step 2: Execute tpsUtil

Step 3: "Select an operation" is highlighted under Operation and the feature "Build tps file" is selected from drop-down list.

Step 4: "Input" is designated under input path or directory, the specimen directory is browsed and any one of the images is selected. It'll display the path of the directory as "Data file=?"

Step 5: "Output", under output file is selected, and a file name with ".tps" extension is entered and saved. It will display the path of saved directory as "Output file=?"

Step 6: To build the tps file: "compute" under Actions, is executed. A list of images will appear, which needs to be check/uncheck as required. Clicking "create" will create the .tps file and "close" to exit tpsUtil.

A screen shot of the executed .tps file is as below:

```
LM=0
IMAGE=Baboon.jpg
LM=0
IMAGE=Chimpanzee.jpg
LM=0
IMAGE=Common Marmoset.jpg
LM=0
IMAGE=Gibbon.jpg
LM=0
IMAGE=Gorilla.jpg
LM=0
```

As the landmarks are yet to be selected, the file records selected landmarks as zero.

3.6.2 <u>TpsDig</u> Tps Digitizer, another component of the Tps Suite software package is the one used to place landmarks on the sample images and saved the digitized data in tps file format.

Phases of operation, involved in the use of tpsDig are illustrated below:

Step 1: TpsDig software package is executed.

Step 2: To open the target .tps file: *File -> Input source -> File...*

Step 3: The Landmarks are placed by selecting the cross-hair icon from the menu bar. By left-clicking we can place the landmarks at the required locations and can delete the selected landmark by right clicking.

Step 4: The landmark data is saved by *File-> Save data-> Save-> Overwrite...*

A screen shot of the executed .tps file post landmarking:

```
LM=2
153.00000 274.00000
144.00000 275.00000
IMAGE=Baboon.jpg
ID=0

LM=2
122.00000 238.00000
106.00000 237.00000
IMAGE=Chimpanzee.jpg
ID=1
```

3.6.3 **PAeontological STatistics (PAST)** Past is a bundled software which was initially created for Paleontological information investigation yet in a current scenario, it is being constantly utilized by different fields like life science, earth science, engineering and stocks in profit making commerce.
Once the data information file (dimensions set as column-specimens put as rows) is imported, the following geometric analysis are obtained from the PAST:

3.6.3.1 **Procrustes fitting** The specific informations imported is selected/ highlighted.

Transform-> Procrustes 2D/3D.

It is necessary to normalize the obtained Landmark data as we need to analyze only the silhouette. Landmarks reduce the dimension impact on the specimens, and so institutionalizing to Procrustes scale we evacuate the rotation, scaling, and translation possessions. Procrustes residual can be obtained by further modification, through deducting the mean from Procrustes; *Transform-> Subtract mean*.

3.6.3.2. **Principal Component Analysis (PCA)**

Multivar-> Principal components

This is a technique exploited to discover the variable called the Principal Component that embodies the greatest difference in the information set. It diminishes the measurements to just the primary constituents. These first-hand variables in terms of a correlation matrix. On the off chance that variable are in indistinguishable units, then variance-covariance technique is put into execution, else it is opt for correlation. The eigenvalues provides a degree of change consistent to the corresponding vector. The rate of change is additionally presented in conjunction.

Succeeding components are assessed and presented in PCA:

3.6.3.2. (A) <u>Jolliffe Cut-off</u> It provides a threshold for the number of components that can be engaged into calculation. Eigenvalues obtained with magnitude below the Jolliffe Cut-off can be discarded as unimportant.

3.6.3.2. (B) <u>Scree plot</u> It's a graphical representation of significant components that can be involved in the analysis. Once curve starts to flatten, the components beyond the lower plateau may be considered as insignificant.

3.6.3.2. (C) <u>Scatter Plot</u> It illustrates all the information focuses strategized in the Cartesian coordinates fixated around two principle components. "Spanning Tree" can additionally be contrived which is the briefest conceivable set of joined lines joining the data centers. It may be utilized to assemble close features concentrated on the Euclidean difference. "Bi-plot option" indicates a projection of the first set of axis onto scatter plot. "View Loadings" demonstrates the gradation of unique features in the distinctive segments.

3.6.3.3. <u>Thin-plate splines and warps</u>

Geomet-> Thin plate splines and warps 2D.

The information is orchestrated in lines and sections. "Expansion factors" alternative shows the magnitude of extension and compression around every data point in yellow numbers which suggests the growth in local domain. The extension demonstrated by this component is coded in color for all the network components where green is characteristic for development and purple for withdrawal of the element feature. Principle strain can likewise be shown at every historic point with significant strain set colored in black and minor strain labeled as brown. These vectors demonstrate directional extension.

3.6.3.3 (A) <u>Partial warps</u> It could be chosen from thin-plate spline window to see the fractional distortion for a specific spline warping. Higher order partial warp indicates an enhanced level of local distortion while first order warping induce global deformation. The projection portion of the twist remains for direct interpretation, scaling, pivot and shearing. At the point when the magnitude of the components is amplified from null to some value, the first milestone setup and a framework will in this way distort the grid as indicated by the specified partial warp.

3.6.3.3 (B) Relative Warps

Geomet-> Relative warp 2D.

Thin-plate transformations of the principal components are called Relative warp. These distort the set from the mean shape to for each of the sample shapes. It offers a selection to obtain uninterrupted PCA of the landmarks. Alpha can be set to any of the three values:

Alpha= -1
It emphasizes local deformation/variation.

Alpha= 0:
This is the PCA of landmarks directly. It is equivalent to the shape PCA but excluding the affine component.

Alpha= 1:
It emphasizes the global deformation/variation.

Importance is the criteria the relative warps are arranged on. The couple of front warps are usually the most enlightening. Thin-plate spline alteration lattices are used to analyze the relative warps. At the point when the bounty factor is expanded or diminished from null, selective relative warp administers the formation of the original landmark configuration and grid distorts according to it.

4. Results and Discussion

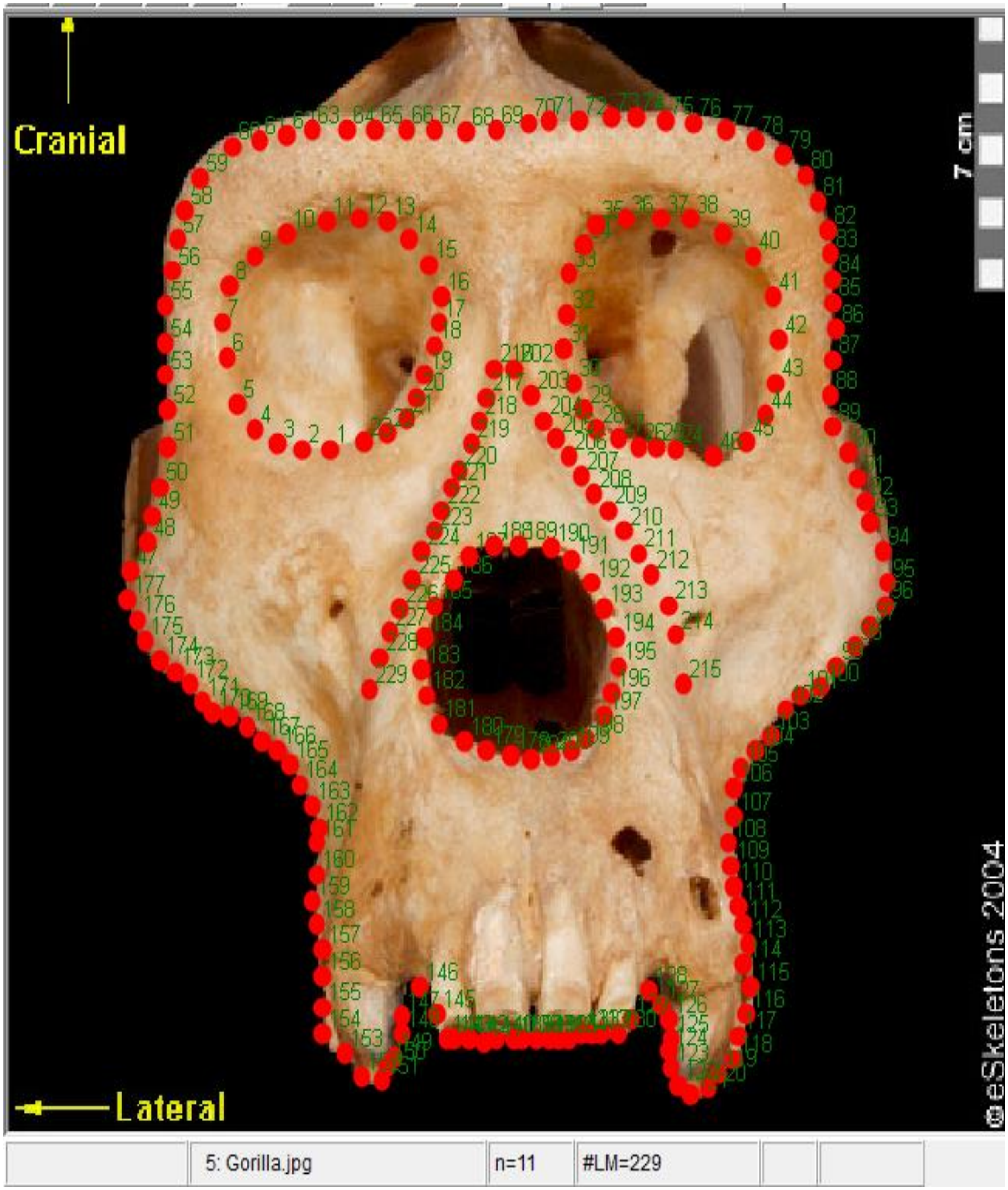


Figure 7: A skull sample (Gorilla) with 229 specified datapoints, is illustrated.

Lateral.TPS

File Edit Transform Plot Statistics Multivar Model Diversity Time Geomet Strat Cladistics Script

Edit mode Edit labels Square mode

	X1	Y1	X2	Y2	X3	Y3	X4	Y4	X5	Y5	X6
0	34.00000	308.00000	38.00000	310.00000	43.00000	312.00000	48.00000	314.00000	52.00000	316.00000	58.00
1	72.00000	142.00000	70.00000	144.00000	66.00000	146.00000	61.00000	149.00000	58.00000	152.00000	51.00
2	310.00000	236.00000	307.00000	239.00000	303.00000	240.00000	300.00000	242.00000	295.00000	243.00000	288.00
3	278.00000	306.00000	273.00000	307.00000	53.00000	173.00000	48.00000	174.00000	43.00000	179.00000	38.00
4	81.00000	336.00000	86.00000	334.00000	89.00000	334.00000	92.00000	334.00000	96.00000	336.00000	100.00
5	370.00000	205.00000	368.00000	202.00000	366.00000	194.00000	365.00000	189.00000	366.00000	198.00000	364.00
6	43.00000	166.00000	41.00000	169.00000	38.00000	172.00000	36.00000	174.00000	33.00000	178.00000	30.00
7	77.00000	359.00000	73.00000	357.00000	70.00000	353.00000	66.00000	350.00000	62.00000	347.00000	58.00
8	385.00000	173.00000	387.00000	168.00000	391.00000	165.00000	394.00000	163.00000	397.00000	157.00000	396.00
9	304.00000	274.00000	301.00000	277.00000	298.00000	279.00000	295.00000	282.00000	292.00000	284.00000	289.00
10	157.00000	284.00000	161.00000	286.00000	164.00000	289.00000	167.00000	291.00000	171.00000	293.00000	174.00

Figure 8: PAST loaded with data for each 11 specimens

- Images are induced with landmarks using tpsDig as illustrated in fig. 7. The data set is subsequently loaded in the software with specimens placed in rows and coordinates set in columns shown in fig. 8.

Principal Components

PC	Eigenvalue	% variance
1	0.0423911	55.667
2	0.0118075	15.505
3	0.00601372	7.897
4	0.00452579	5.9431
5	0.00329636	4.3287
6	0.00248271	3.2602
7	0.0018911	2.4833
8	0.00148747	1.9533
9	0.0014415	1.8929
10	0.000814496	1.0696

Table 1: Principal Components are presented with % variance and eigenvalue

- Once the general Procrustes Alignment has been achieved, Principal Component Analysis is undertaken. This process reduces the dimensions required to include in the analysis and provides the significant Principal components.

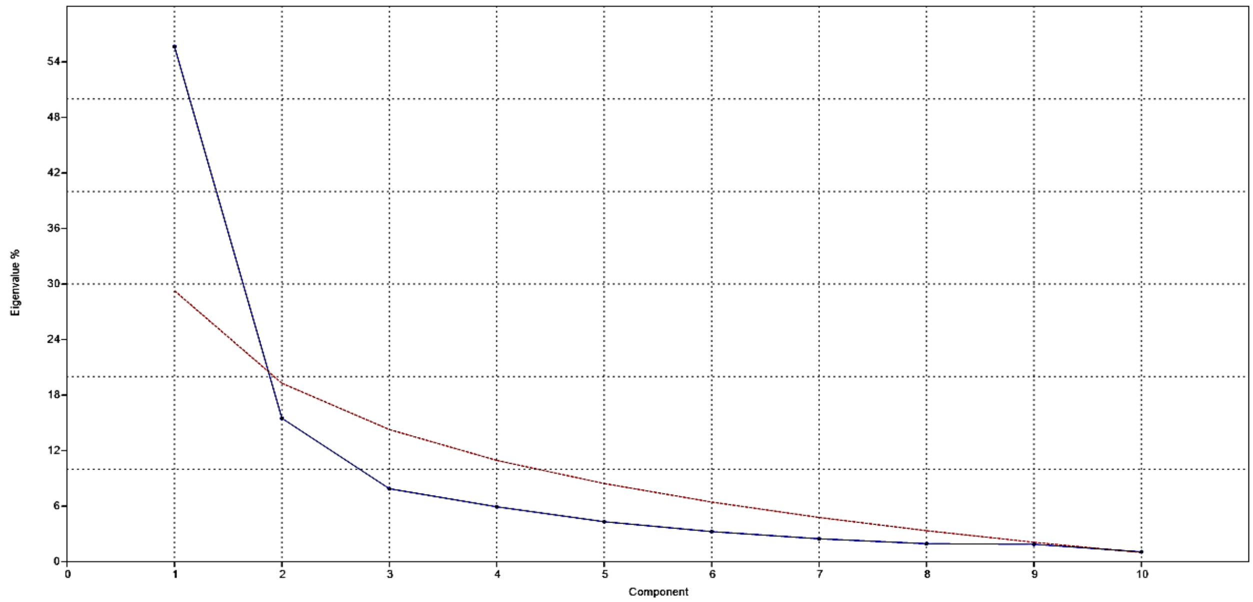


Figure 9: scree plot

- Scree plot is a significant scheme which displays that the grade dies after 10 components so 11 extents are abridged to 10.

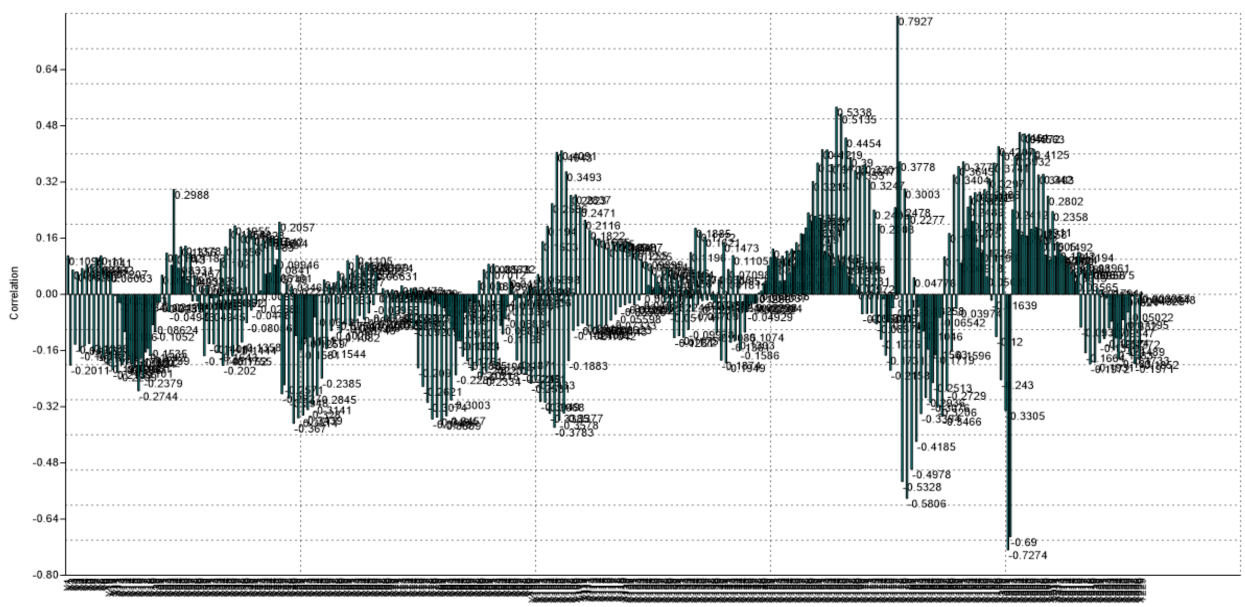


Figure 10: Component loadings

- Loadings stratagem indicates the step of the different unique features present in the diverse modules.

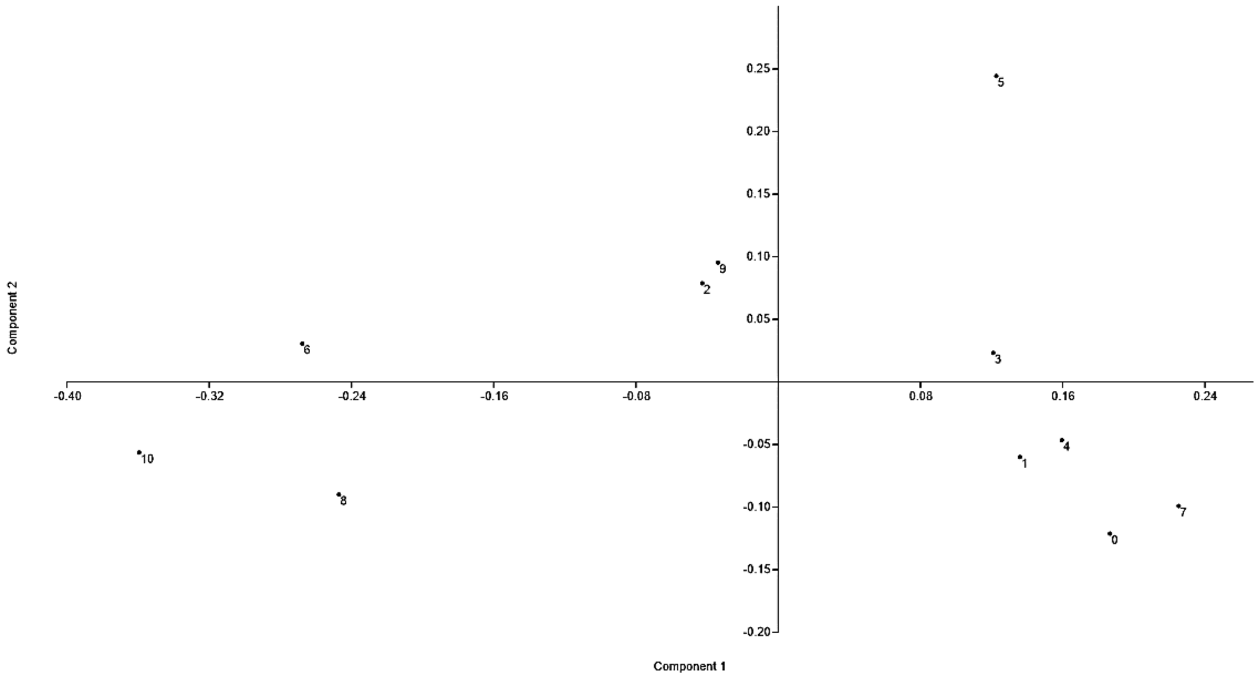


Figure 11: Scatter plot

- Scatter design is one of the greatest revealing device. Correlation of the species are illustrated in terms of shape.

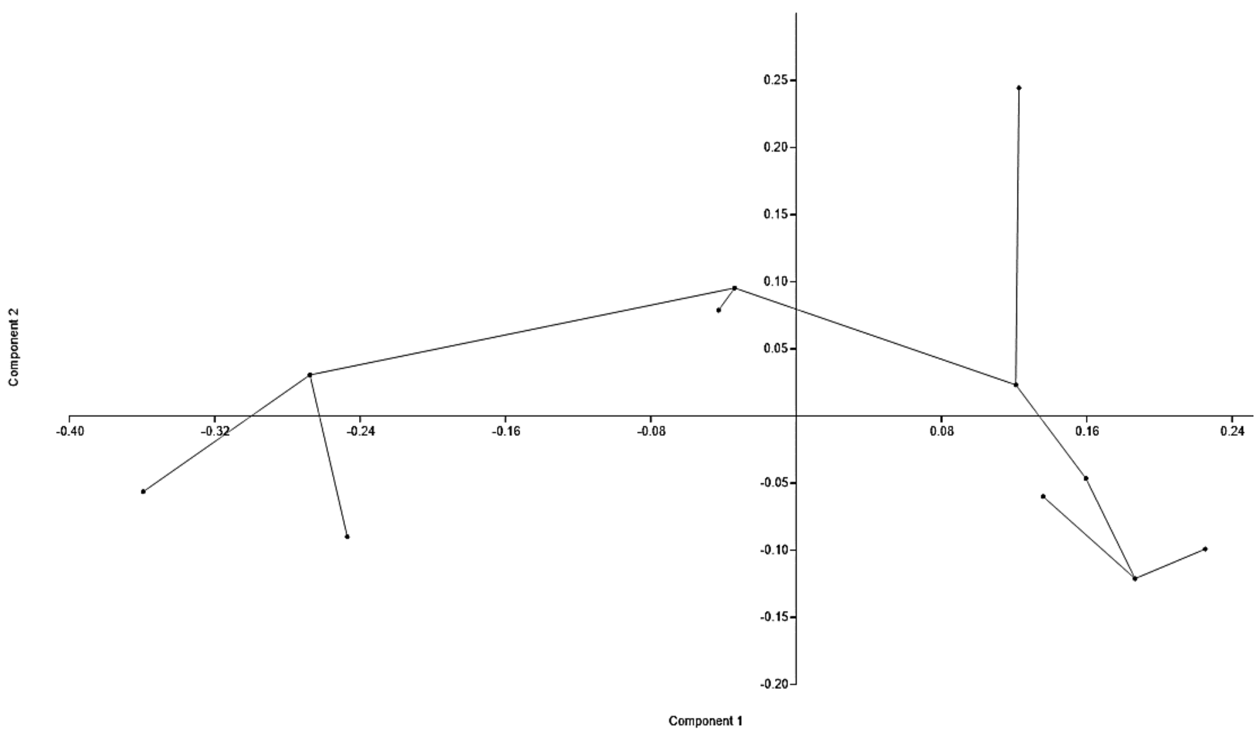


Figure 12: PCA minimal spanning tree

- The minimal spanning tree connects each component with its closest corresponding element. This plot enables the mathematical results obtained in this study to be associated with the established anthropological evolutionary pathway, implying an accurateness of this investigation.

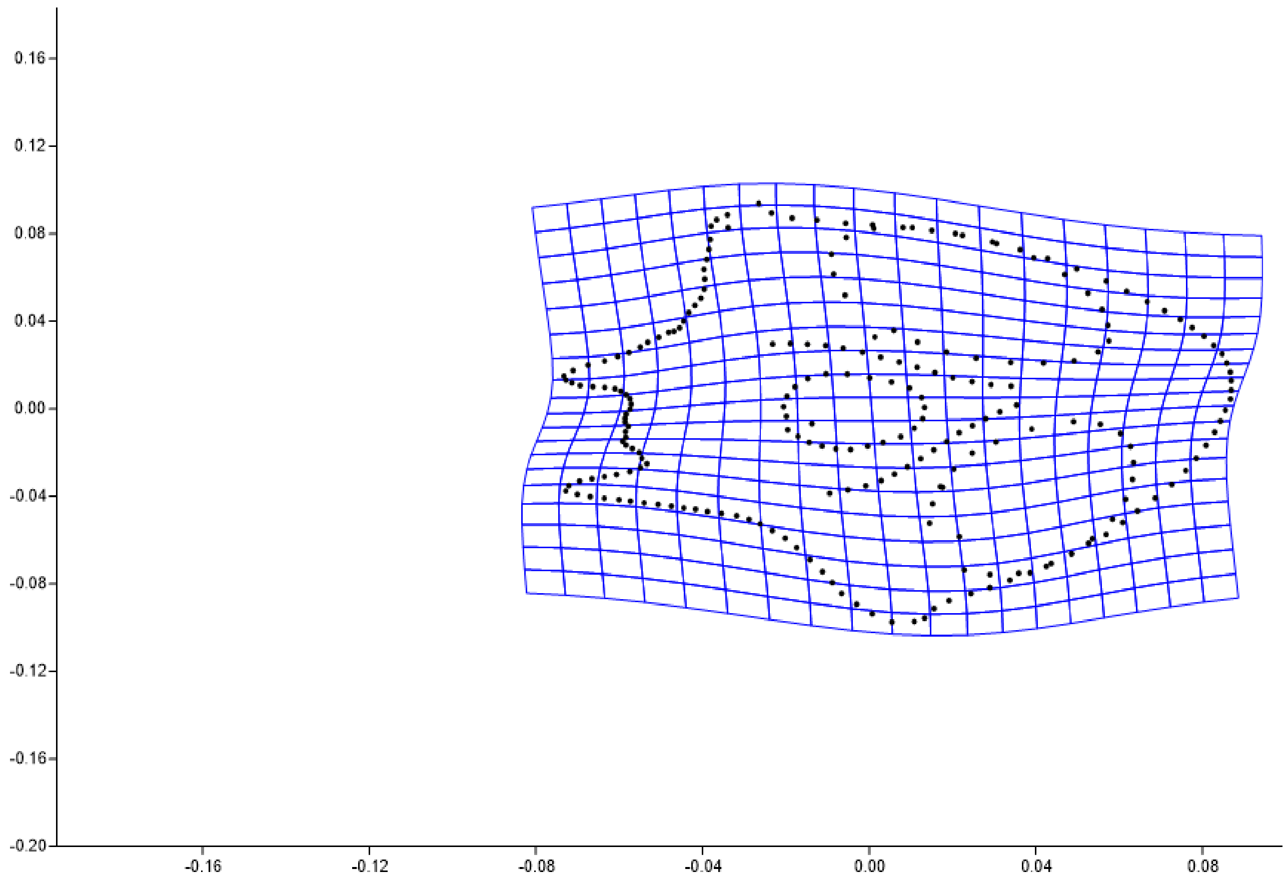


Figure 13: Partial warp 4 with amplitude 3.

- The twist of warp 4 with amplitude 3 is illustrated. This analysis is carried out exclusively along the ventral domain, so as to demonstrate the possible deformations of the particular specimens, along the feature rich plane. All the probable distortion of specific samples can be perceived and related.
- Relative distortion as in PCA plots cases in a area of two utmost significant components of TPS like in PCA.

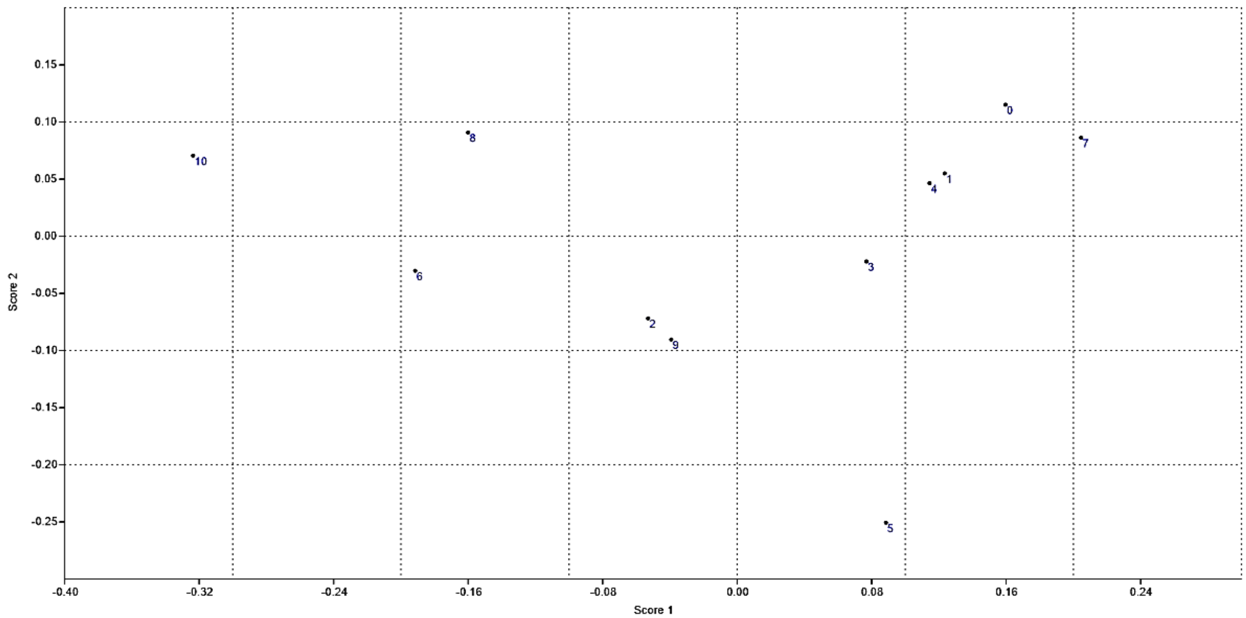


Figure 14: Relative Warp plot

➔ Analysis in the dorsal plane:

- The Jolliffe cutoff observed in this analysis is 0.04971, implying this study has undertaken 6 principal components that are significant in their Eigen values.
- The principle components PC 1 and PC 2 are observed to have returned % variance of 23.442 and 18.996, respectively.
- PCA scatter indicates that the Specimen 1 (Baboon) and 10 (Tarsier) are present at an equivalent detachment from PC1 implying that they have resemblance in the cranium shape. While specimen 2 (Common Marmosets) and 6 (Lesser Bushbaby) show comparable shape conformation. Specimen 7 (Orangutan) is quite close to 8 (Ruffed Lemur). Specimen 5 (Human) and specimen 9 (Squirrel monkey) is found to be fairly unlike all others.

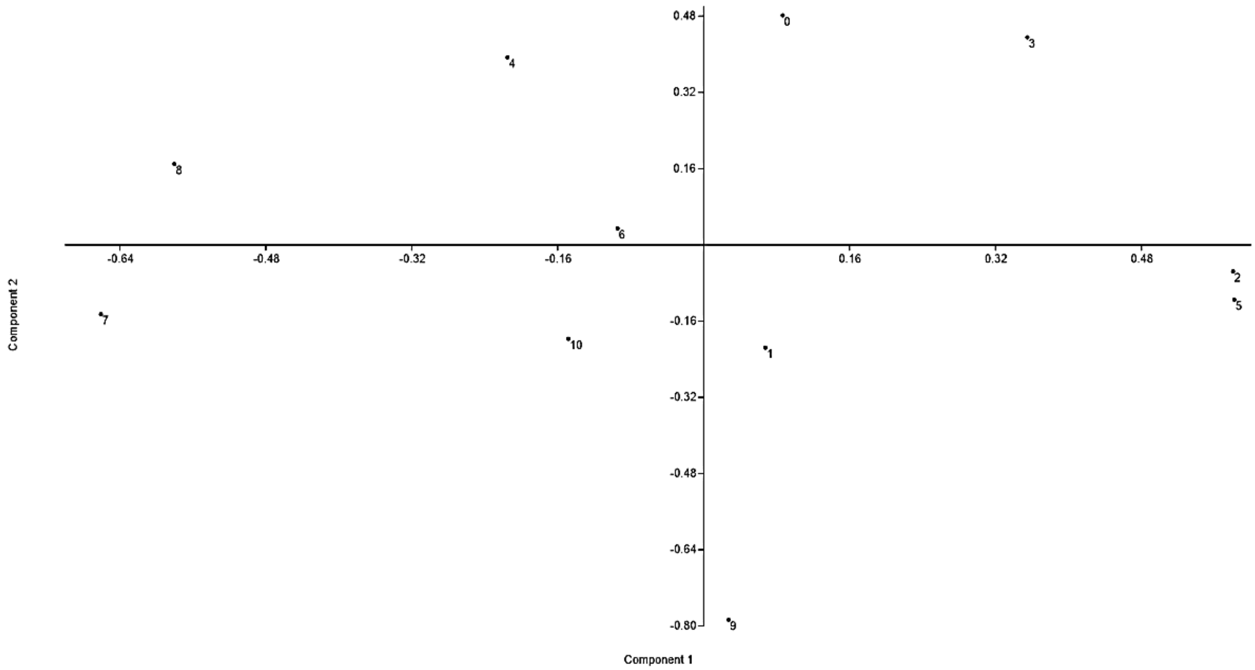


Figure 15: PCA scatter plot for dorsal plane analysis

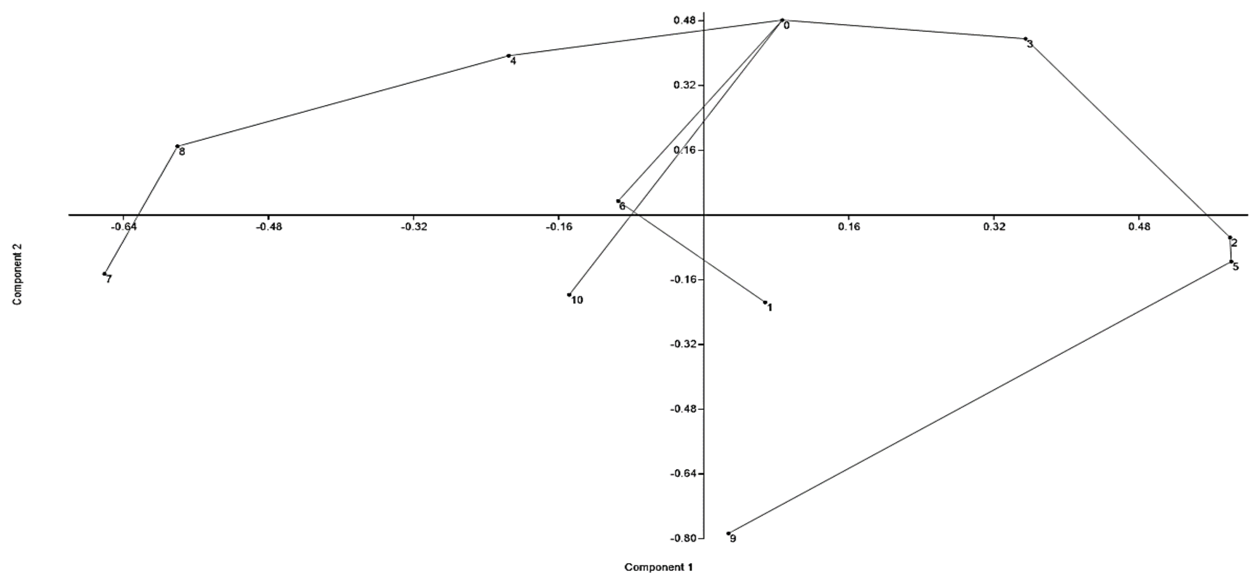


Figure 16: Minimal spanning tree for dorsal plane analysis

- The minimal spanning tree obtained from the dorsal plane analysis is similar to the already established anthropological evolutionary hierarchy, indicating a correct trend in the analysis.

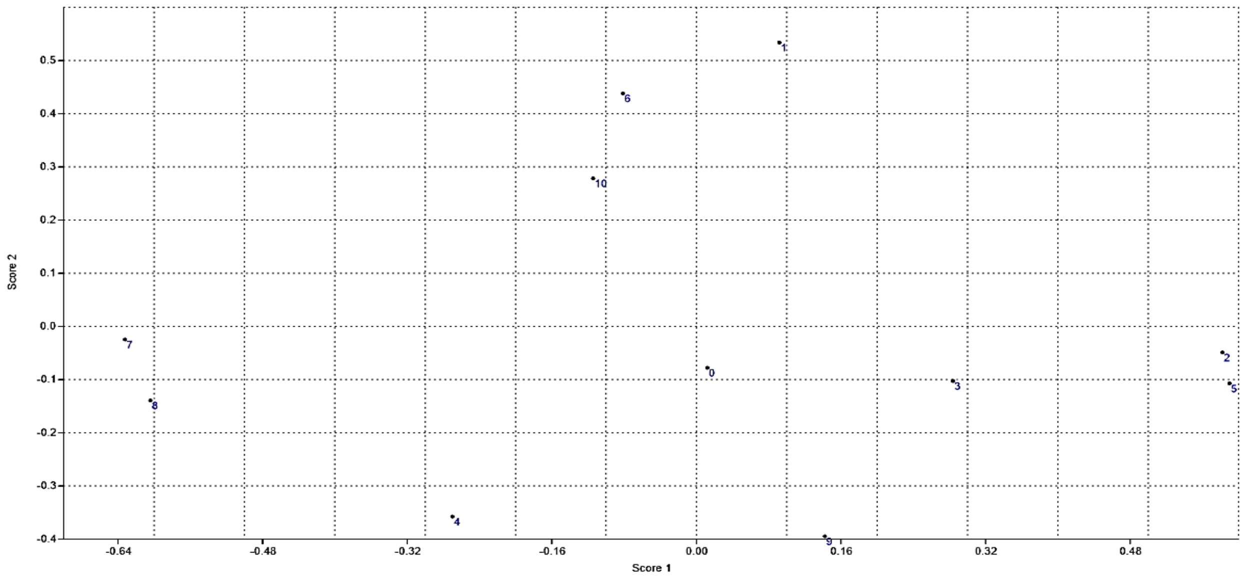


Figure 17: Relative warp plot for dorsal plane analysis

- Relative warp as in PCA, illustrates the samples in a space of dual significant constituents. This shows that specimen 6 (Lesser Bushbaby) has parallel distortion to specimen 10 (Tarsier), though 1 (Chimpanzee) is also near. 3 (Gibbon) and 4 (Gorilla) are far apart. 7 (Orangutan) is near to 8 (Ruffed lemur) and 2 (common marmoset) is in close proximity to sample 5 (Homo sapiens).

➔ Analysis in the Lateral plane:

- The Jolliffe cutoff observed in this analysis is 0.03867, implying this study has undertaken 6 principal components that are significant in their Eigen values.

- The principle components PC 1 and PC 2 are observed to have returned % variance of 20.452 and 17.096, respectively.

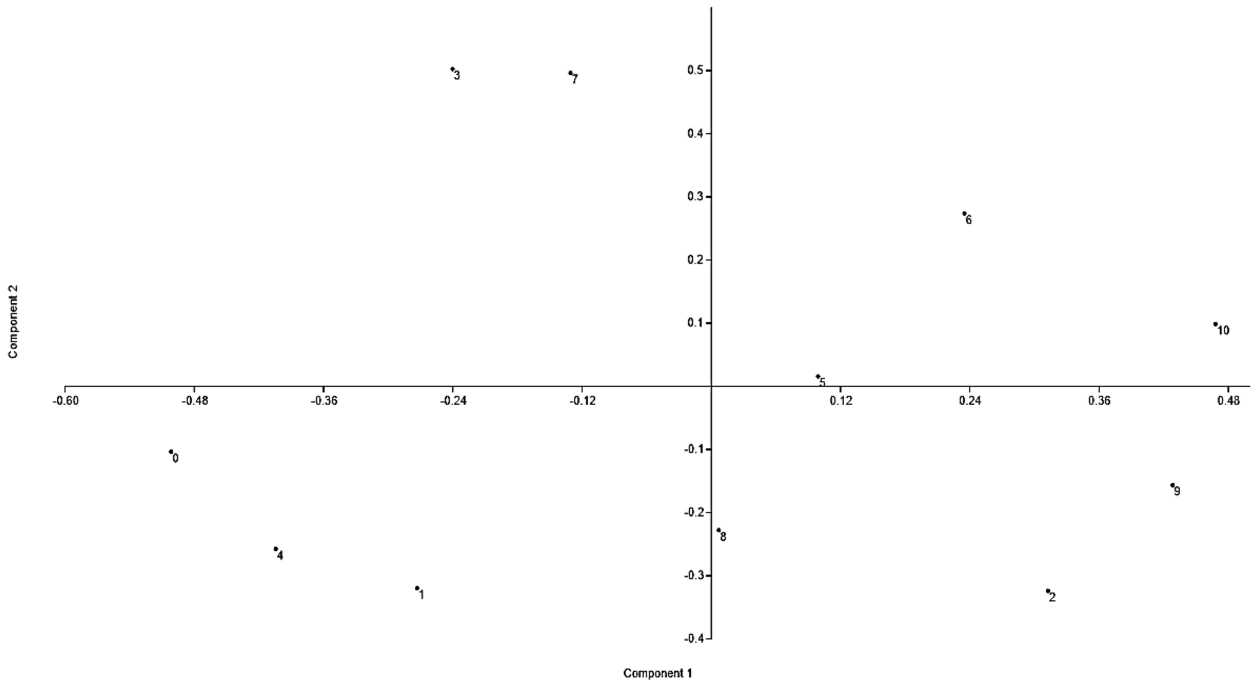


Figure 18: PCA scatter plot for Lateral plane analysis

- The PCA scatter in this investigation illustrated that Specimen 3 (Gibbon) and 7 (Orangutan) portion the identical remoteness from Principle Component 1 resources they obligate resemblance in skull shape along this visual axis. Specimen 0 (Baboon) and 10 (Tarsier) show comparable shapes. Specimen 2 (Common Marmoset) is quite close to 9 (Squirrel monkey) and 1 (Chimpanzee) is in close proximity to sample 4(Gorilla). Specimen 6 (Lesser Bushbaby) is found to me quite different from all others.
- The Homo sapiens skull, i.e. specimen 5 is in close conjugation with the principle component axes, indicating it to be on the verge of a mean shape in this relative comparison.

- The minimal spanning tree in insignificant in the analysis of the lateral plane as the development of the facial features in terms of the muzzle length, the protruding of the snout, or the characteristics “flat face” feature found on the higher order primates is significantly dependent on the individual specimen’s food habits, rather than on the environmental conditions. Carnivores tend to have a longer facial profile when observed normally towards the lateral plane. They have well developed canine dental structure so as to be efficient in their feeding process. Individual species that are dependent on the vegetation, fauna, and plant products incline towards having a flatter face profile.
- The relative warp plot shows some dissimilar outcomes from PCA like sample 6 (Lesser Bushbaby) holds comparable distortion with example 9 (Squirrel Monkey) though 10 (Tarsier) is also near. 2 (Common marmoset) and 8 (Ruffed Lemur) are far apart. But 1 (Chimpanzee) is near to 4 (Gorilla) and sample 3 (Gibbon) is close to specimen 7 (Orangutan) which is similar as in PCA.

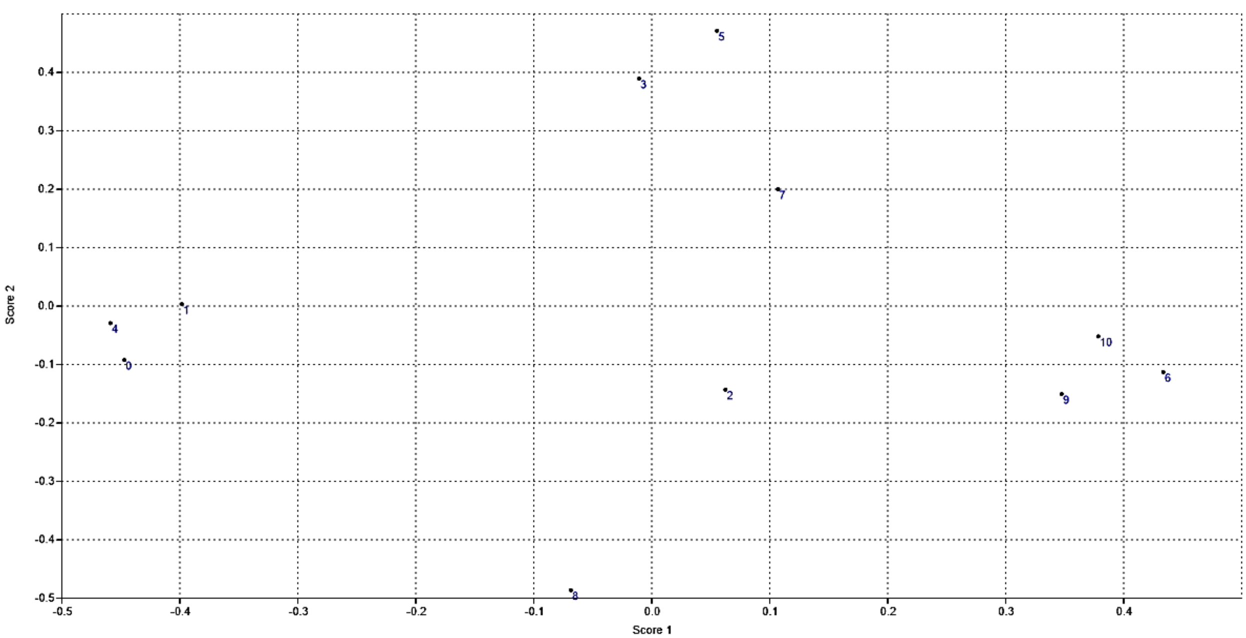


Figure 19: Relative warp plot for lateral plane analysis

➔ Analysis in the Ventral plane:

- The ventral plane is the most important domain in this investigation, as it contains the maximum number of features that have been included in this study and has effectively made it feasible to compare various primates over the evolutionary pathway.
- The Jolliffe cutoff observed in this analysis is 0.005331, implying this study has undertaken 3 principal components that are significant in their Eigen values.
- The principle components inclusive in this investigation; PC 1, PC 2 and PC 3 have been observed to have returned % variance of 55.667, 15.505 and 7.897 respectively.
- This is this high value of % Variance in the primary principal components that makes the investigation in this plane and its subsequent results a pivot to the success of this research.
- The PCA scatter in this investigation demonstrated that Specimen 3 (Gibbon) and 6 (Lesser Bushbaby) share the equal distance from Principle Component 1 implying that the samples have characteristics resemblance in skull shape. Specimen 1 (Chimpanzee) and 10 (Tarsier) appears to have comparable shapes. Specimen 2 (Common Marmoset) is quite close to 9 (Squirrel monkey).Specimen 1 (Chimpanzee) is in close proximity to sample 4(Gorilla), and specimen 0 (Baboon) is grouped together with 7 (Orangutan).

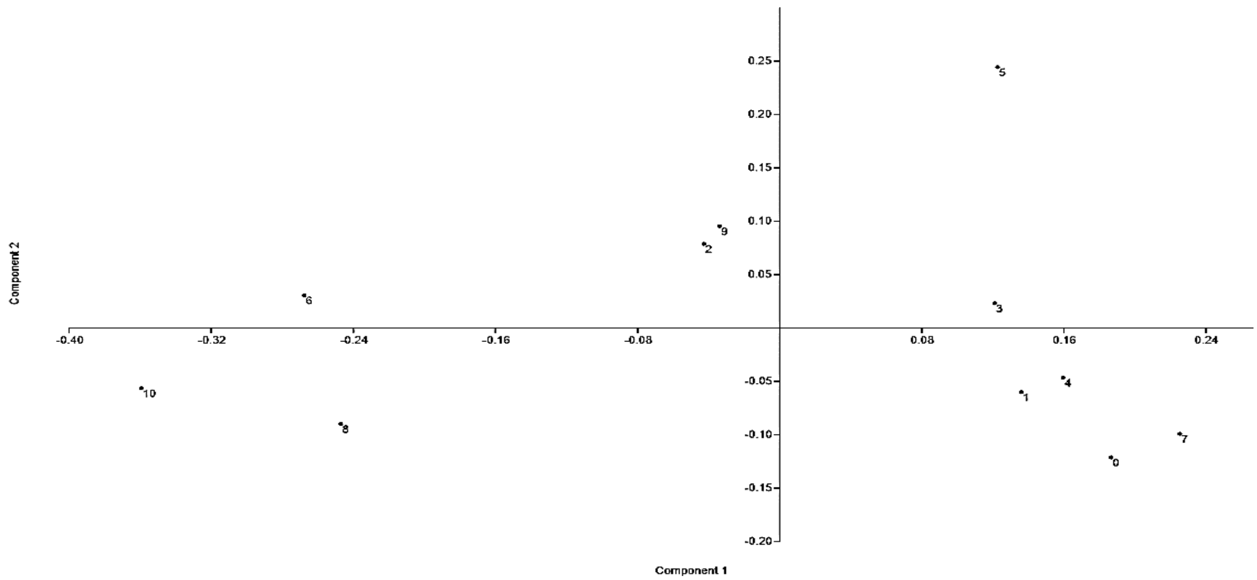


Figure 20: PCA scatter plot for Ventral plane analysis

- Specimen 5 (Homo sapiens) is found to be quite different from all others.
- The minimal span tree obtained by virtue of the analysis along the ventral plane is an accurate recreation of the “already established” evolutionary hierarchy. This also acts as an indigenous endorsement of the accuracy of this investigation and the reliability and repeatability of the results obtained.

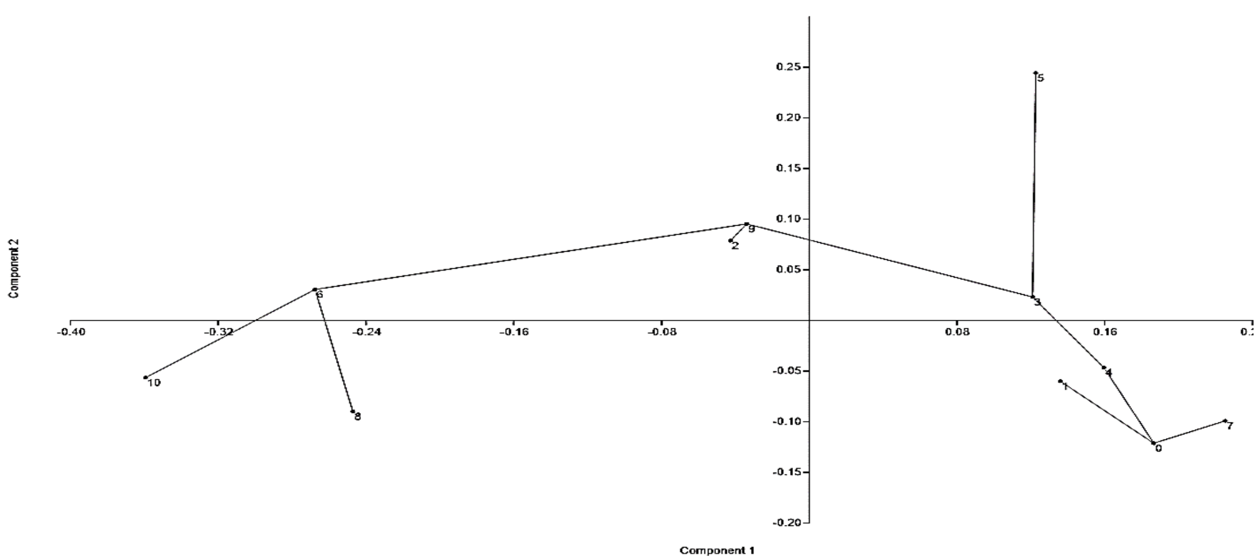


Figure 21: Minimal spanning tree for Ventral plane analysis

- The minimal span tree obtained in this study has four observable nodes that might resemble evolutionary branching in the discipline of anthropology.
- The first diversification is observed at sample 6 (lesser Bushbaby), which give rise to two terminal species, 8 (Ruffed Lemur) and 10 (Tarsier). Specimen 9(Squirrel Monkey), further evolved from sample 6, branches out to 2(Common Marmoset). Specimen 3 (Gibbon) evolves into 4 (Gorilla) and 5(Human- Homo sapiens). Species 4, evolves into 0 (Baboon), which branches out to form 1 (Chimpanzee) and 7(Orangutan).
- A comprehensive illustration of the above discussed evolutionary development is compiled in Figure 22.

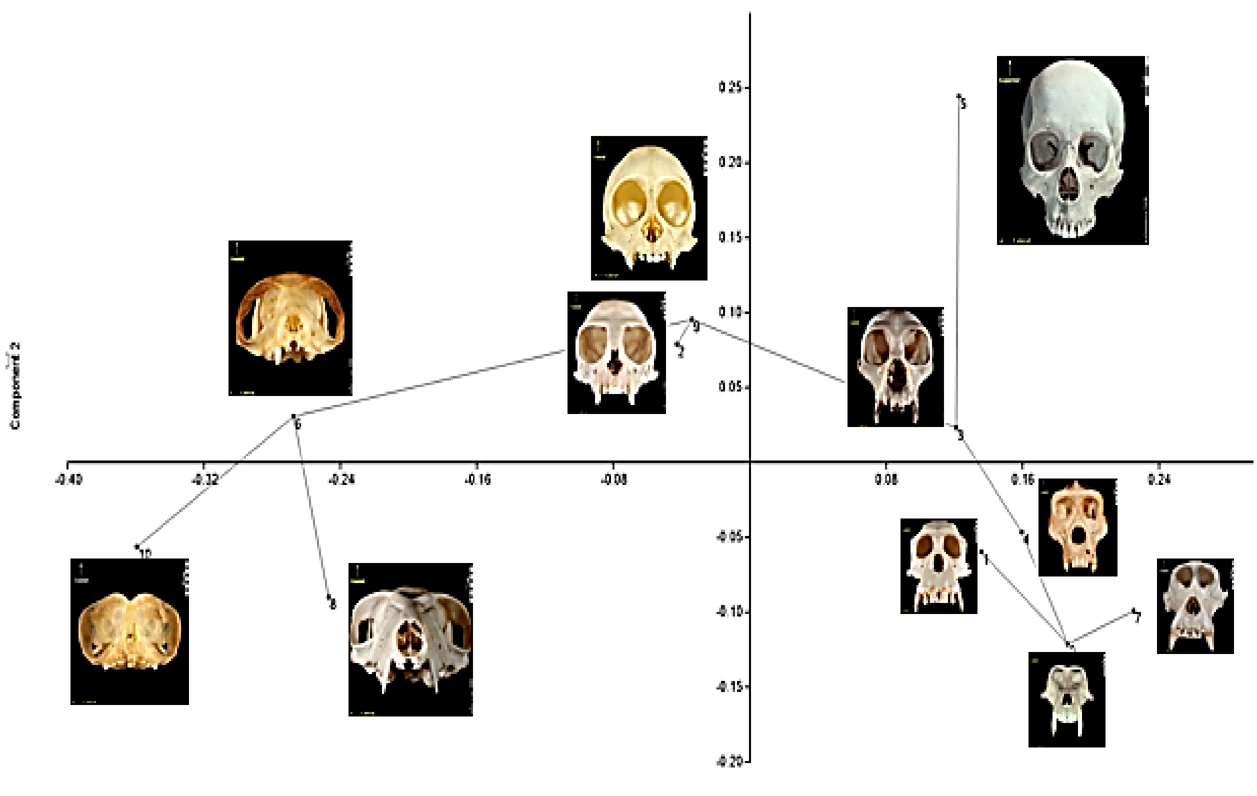


Figure 22: Correlating Evolutionary development with the Minimal span tree.

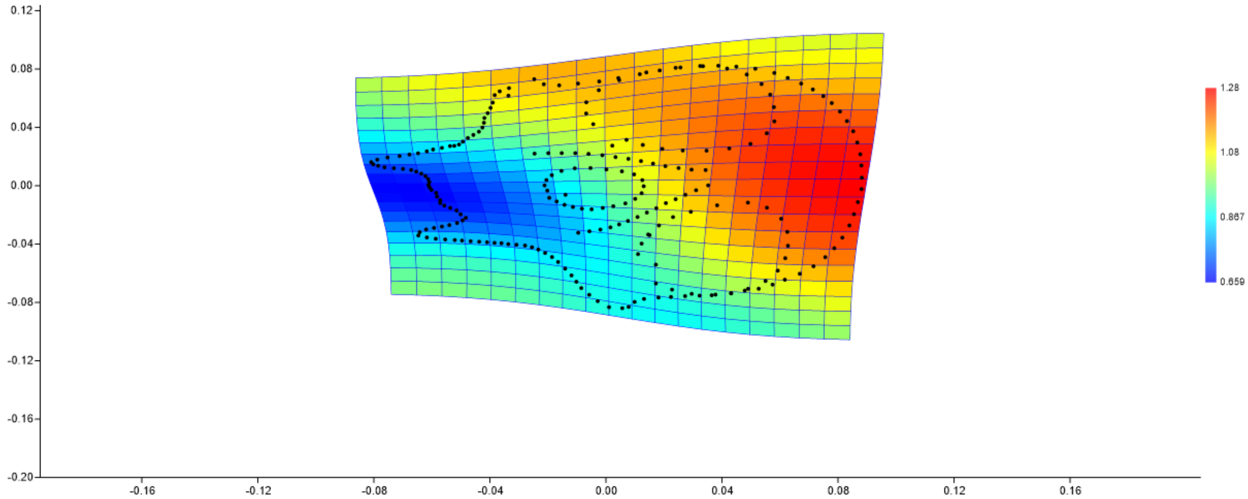


Figure 23: Partial warp

- Partial Warps illustrates the distortion of warp 2 with heft 3. Likewise, all the possible distortion of any particular specimens can be perceived and equated. It has been found that 2 (common marmoset) and 9 (Squirrel Monkey) specimens have similar deformation while Sample 0 (Baboon), 1 (Chimpanzee), 3 (Gibbon), 4 (Gorilla) and 7 (Orangutan) have similarity in the obtained distortion. In the same way, 6 (Lesser Bushbaby) possesses similar deformation with respect to specimen 8 (Ruffed Lemur).

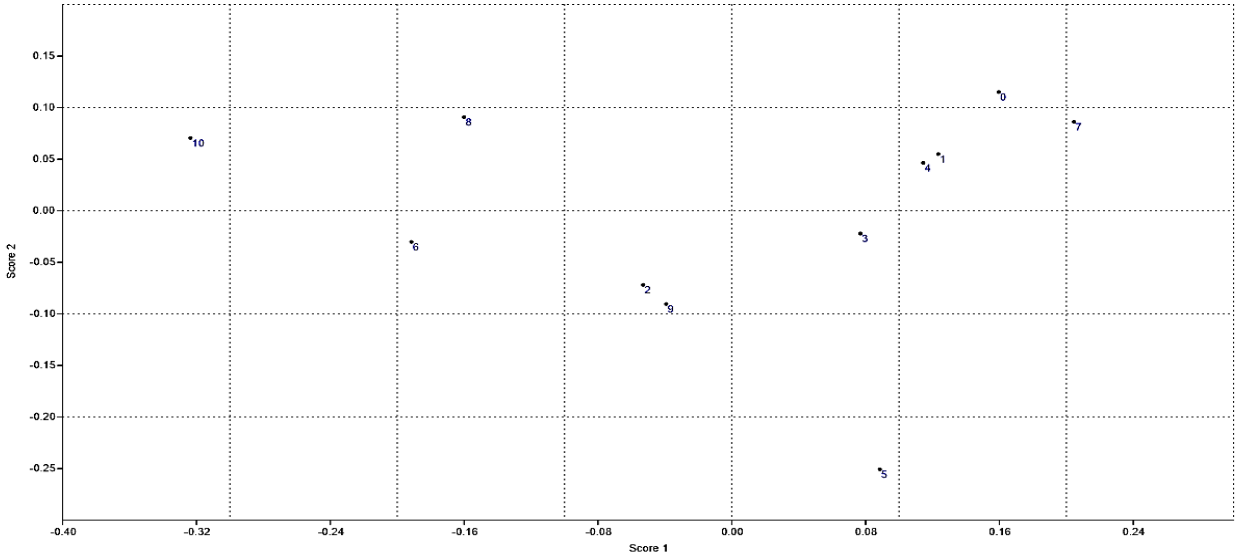


Figure 24: Relative warp plot for ventral plane analysis

- This plot in ventral plane analysis returns results identical to the ones obtained from PCA like specimen 6 (Lesser Bushbaby) possesses similar deformation with respect to specimen 8 (Ruffed Lemur) and model 11 (Tarsier) is also adjacent. Specimen 5 (Homo sapiens) is spaced far out. Sample 0 (Baboon), 1 (Chimpanzee), 3 (Gibbon), 4 (Gorilla) and 7 (Orangutan) are grouped together and the specimen 2 (common marmoset) and 9 (Squirrel Monkey) are in close proximity to each other.
- The result obtained from this relative warp plot is in agreement with the output returned by the Principle component scatter analysis. This proves a linearity in this undertaken study and a reliability of the obtained results.

5. Conclusion and Outlook

- **Conclusion:** The primary objective of this research was to ascertain the similarities and the differences between the cranium structures of various primates based on a geometric morphometric approach. Eleven specimens were obtained and the study was undertaken in there different planes, to impart a three-dimensionality characteristic in the investigation, termed as "Pseudo 3D" analysis. 435 landmarks were used as regions of interest while analyzing along the dorsal plane, 290 landmarks were initiated in each of the lateral domain samples and 229 landmarks were used to digitize the samples obtained in the ventral domain. Principal component analysis was successfully implemented. The Jolliffe cutoff was calculated to be 0.04971, 0.03867and 0.005331in dorsal, lateral and ventral samples. Lower the Jolliffe cutoff, higher is the number of features included in the analysis. Scree plot illustrated efficient decline of aspect from 11 to 10 in PCA analysis. The utmost significant results obtained are that samples *0(Baboon)*, 1 (*Chimpanzee*), 3 (*Gibbon*), 4 (*Gorilla*) and 7 (*Orangutan*) are grouped together, while 6 (*Lesser Bushbaby*), 8 (*Ruffed Lemur*) and 10 (*Tarsier*) form a group of closely related species and the "rodent" like primates of 2 (*Common Marmoset*) and 9 (*Squirrel Monkey*) are in close proximity to each other. *Homo sapiens* specimen tends to be different from all other species due to variation in characteristics, and standing on top of the Suborder Anthropoidea, Infraorder Catarrhini evolutionary pyramid. The *Minimum span trees* obtained from our analysis is identical to the established evolutionary hierarchy illustrating the accuracy of this investigation. Therefore, the objective of the investigation is efficaciously realized.

- **Outlook:** The minimum span tree obtained from the analysis in the dorsal and lateral planes do not coincide with the evolutionary setup. This can be due to the reason that the mechanism of evolution of the ventral cranium features is different from the lateral and dorsal components. Critics can effectively induce that the only reason the analysis along the ventral obtains a minimal span tree, i.e. conjugated with the established evolutionary pathway is because of an abundance of features available to the ventral feature space unlike the lateral and dorsal planes that have a lack of conventional characteristics. Additionally, the lack of exhaustive GM work based on lateral and dorsal plane, makes it more difficult to correlate the obtained observations with previous investigation. We did obtain some relatively optimistic results but further studies in this discipline as well as expansion to various other body parts, based on versatile comparison is highly essential. Due to the immense diversity of fauna along the world tropical belt, this may be the finest place for similar morphometrics exploration. The further studies that needs to be undertaken, has to be extended from this "Pseudo 3D" domain to a real time three dimensional space. Additionally, the above results are obtained on the basis of analysis on cranium images, but images provide information segregated in the two- dimensional domains of the physic-visual Planes. Hence, extensive further study is required in this domain.

6. References

1. D.C. Adams, F.J. Rohlf, and D.E. Slice. *Geometric morphometrics: ten years of progress following the 'revolution'*. Italian Journal of Zoology, 71(1):5-16, 2004.
2. D. Adriaens. Workshop *"deformities in Fish larvae". In Geometric morphometrics as a useful tool for visualizing and analyzing deformities in Fish*, March 11, 2005.
3. L. Rüber and D. C. Adams. *Evolutionary convergence of body shape and trophic morphology in cichlids from Lake Tanganyika*. Journal of Evolutionary Biology, 14:325-332(8), March 2001.
4. F. James Rohlf. Morphometrics [*http://life.bio.sunysb.edu/morph/*], 02/2010.
5. Bookstein, F. L. *A hundred years of morphometrics*. Acta Zoologica Academiae Scientarium Hungaricae, 44(1–2), 7–59, 1998.
6. Rohlf, F. J., & Marcus, L. F. *A revolution in morphometrics. Trends in Ecology & Evolution*, 8(4), 129–132, 1993
7. O'Higgins, P. *The study of morphological variation in the hominid fossil record: Biology, landmarks and geometry*. Journal of Anatomy, 197(1), 103–120, 2000.
8. Gunz, P., Mitteroecker, P., & Bookstein, F. L. *Semi landmarks in three dimensions. In D. E. Slice (Ed.), Modern morphometrics in physical anthropology*. New York: Kluwer Academic/Plenum Publishers, 2005
9. Bookstein, F. L. *Biometrics, biomathematics and the morphometric synthesis. Bulletin of Mathematical Biology*, 58(2), 313–365, 1996.
10. Small, C. *The statistical theory of shape*. New York: Springer, 1996

11. Dryden, I. L., & Mardia, K. V. *Statistical Shape Analysis. New York: John Wiley and Sons*, 1998.

12. Lele, S., & Richtsmeier, J. T. *Euclidean distance matrix analysis: A coordinate free approach for comparing biological shapes using landmark data*. *American Journal of Physical Anthropology*, 86, 415–428, 1991.

13. Lele, S., & Richtsmeier, J. T. *An invariant approach to statistical analysis of shapes*. Boca Raton, Fla: Chapman & Hall/ CRC, 2001.

14. Lestrel, P. *A Fourier analytic procedure to describe complex morphological shape. In A. Dixon & B. Sarnat (Eds.), Factors and mechanisms influencing bone growth* (pp. 393–409). Los Angeles: Alan R. Liss, 1982

15. Ashburner, J., & Friston, K. J. *Voxel-based morphometry— The methods. NeuroImage*, 11(6 I), 805–821, 2000.

16. Blackith, R. E., & Reyment, R. A. *Multivariate morphometrics.* London, New York: Academic Press, 1971.

17. Marcus, L. F. *Traditional morphometrics*, 1990.

18. In F. J. Rohlf & F. L. Bookstein (Eds.), *Proc. Michigan Morphometrics Workshop* (pp. 77–122). Ann Arbor, Michigan: Univ. Michigan Museums.

19. Oxnard, C. E. *The order of man: a bio-mathematical anatomy of the primates*. Hong Kong: Hong Kong University Press, 1983.

20. Baak, J. P., & Oort, J. *Morphometry in Diagnostic Pathology*. New York: Springer, 1983.

21. Weibel, E. R. (1979). *Stereological Methods: Practical Methods for Biological Morphometry*. New York: Academic Press.

22. Baddeley, A. J., & Vedel Jensen, E. B. *Stereology for Statisticians*. Boca Raton: Chapman & Hall/CRC, 2004.

23. Rohlf, F. J. *Statistical power comparisons among alternative morphometric methods.* American Journal of Physical Anthropology, 111(4), 463–478, 2000a.

24. Rohlf, F. J. *On the use of shape spaces to compare morphometric methods*. Hystrix, 11(1), 9–25, 2000b.

25. Rohlf, F. J. *Bias and error in estimates of mean shape in geometric morphometrics*. Journal of Human Evolution, 44(6), 665–683, 2003.

26. Sensen, C. W., & Hallgrimsson, B. (Eds.). *Advanced Imaging in Biology and Medicine: Technology*, Software Environments, and Applications. Berlin: Springer, 2009.

27. Marcus, M. and Minc, H. *Introduction to Linear Algebra*. New York: Dover, p. 145, 1988.

28. Miriam Leah Zelditch, Donald L. Swiderski, H. David Sheets and Willian L. Fink; *Geometric Morphometrics for Biologists: a primer;*

29. Ovyind Hammer, *Morphometrics-brief notes*, Palaontologisches Institut und Museum, zurich.

30. Karen L. Baab, *The taxonomic implications of cranium shape variation in Homo erectus*. 2007.

31. David F. Wiley, Nina Amenta, Dan A. Alcantara, Deboshmita Ghosh,Yong J. Kil, Eric Delson, Will Harcourt-Smith, F. James Rohlf, Katherine St. John, Bernd Hamann *Evolutionary Morphing*. 2005.

32. Katerina Harvati and Timothy D. Weaver *Human Cranial Anatomy and the Differential Peservation of Population History and Climate Signatures*, 2006.

33. T. Seenivasagan, Kavita R. Sharma, Anchal Shrivastava, B.D. Parashar, S.C. Pant & Shri Prakash *Surface Morphology and morphometric* *analysis of sensillia of Asian tiger mosquito, Aedes albopictus (Skuse): an SEM investigation*. 2009.

34. Elisabeth Nicholson and Katerina Harvati *Quantitative Analysis of Human Mandibular Shape Using Three-Dimensional Geometric Morphometrics*, 2006.

35. F. James Rohlf and Leslie F. Marcus; *A Revolution in Morphometrics*, Vol. 8, No.4..

36. Li Shen, Andrew Saykin, Tara McHugh, John West, Laura Rabin, Heather Wishart, Moo K. Chung, Fillia Makedon. *Morphometric Analysis of 3D Surfaces: Application to Hippocampal Shape in Mild Cognitive Impairment.*

37. Lindsay I. Smith *A tutorial on Principal Components Analysis*, Feb 26, 2002.

38. Heidi Schutz, Jonathan krieger *Guide to geometric morphometrics*.. Version 0.4, May 2007.

39. Oyvind Hammer, *PAST Paleontological Statistics*, Version 2.07 Reference Manual. 1999-2011.

40. Matthias Felix Geiger *A morphometric and molecular study of selected European polpulations of Gymnocephalus cernuus (Linnaeus, 1758)* (Pisces: Teleostei: Percidae), *April 2006.*

41. Image Collection:

- http://www.boneclones.com/catalog-human-adult-skulls.htm
- http://www.channaimuseum.org

7. Abbreviations

1. GPA = Generalized Procrustes Alignment
2. PCA = Principal Component Analysis
3. TPS = Thin Plate Spline
4. L = Landmark
5. PC = Principal Components
6. GM = Geometric Morphometric

8. List of Figures and Tables

Figures:

Tables: